

University of Massachusetts Amherst ScholarWorks@UMass Amherst

Open Access Dissertations

9-2013

The Structure of Consciousness

Lowell Keith Friesen

University of Massachusetts Amherst, lfriesen@philos.umass.edu

Follow this and additional works at: https://scholarworks.umass.edu/open_access_dissertations



Part of the [Other Neuroscience and Neurobiology Commons](#), [Philosophy Commons](#), and the [Psychology Commons](#)

Recommended Citation

Friesen, Lowell Keith, "The Structure of Consciousness" (2013). *Open Access Dissertations*. 794.

<https://doi.org/10.7275/5389-r721> https://scholarworks.umass.edu/open_access_dissertations/794

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

THE STRUCTURE OF CONSCIOUSNESS

A Dissertation Presented

by

LOWELL KEITH FRIESEN

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2013

Philosophy

© Copyright by Lowell Keith Friesen 2013

All Rights Reserved

THE STRUCTURE OF CONSCIOUSNESS

A Dissertation Presented

by

LOWELL KEITH FRIESEN

Approved as to style and content by:

Joseph Levine, Chair

Louise Antony, Member

Hilary Kornblith, Member

Erik Cheries, Member

Hilary Kornblith, Department Chair
Philosophy

To Pam, Lyndon, and Kira.

ACKNOWLEDGMENTS

I would, first of all, like to thank the faculty, students, and support staff at the University of Massachusetts philosophy department for creating the positive study environment I had the privilege of being a part of during my time there. Not all graduate students in philosophy have this good fortune. I did not take my experience for granted and I hope that I made as much of a positive contribution to the department as others did for my benefit. I want to thank my professors for the excellent philosophical education I received. I want to thank Louise Antony, Hilary Kornblith, and Peter Graham for providing me with a positive experience when I was a teaching assistant. I thank Fred Feldman and Beth Grybko for helping me navigate institutional regulations and complete the numerous pieces of paperwork that accompany a graduate degree, especially the additional hurdles I had to jump through as an international student. I thank Phillip Bricker, Peter Graham, Joseph Levine, and my fellow students who were a part of the dissertation seminars during my time at UMass. Thank you for your feedback and your suggestions. They were instrumental in moving my dissertation forward. I thank Heidi Buetow and Jim Binkoski for the wonderful memories I have from the years we spent together in our cramped office. And thank you Scott Hill for the many times you listened to me complain about my dissertation.

I thank Joseph Levine for supervising my dissertation. I valued your feedback on my work and appreciated the practical guidance, the advice, and the occasional nudge you gave me along the way. Thank you for accommodating the inconveniences brought about by my decision to complete my dissertation from Winnipeg and for

being understanding about the familial responsibilities that competed for my attention during this process. I would like to thank the rest of my committee members, Louise Antony, Hilary Kornblith, and Erik Cheries, for reading my work and providing feedback when I requested it.

I would like to thank the many people from outside the UMass community who provided encouragement to me along the way, especially after I moved back to Winnipeg: Chris Banman, Mark Barber, Keith and Tannis Collins, Randall Holm, Roy Jeal, Graham MacFarlane, Greg Monks, John and Shari Pendergrast, Ryan Rempel, Matthew and Betty-Ann Siebert, Karen Sunabacka, David and Ruth Widdicombe, and Jack Zupko.

I would like to thank my extended family. Thank you for always showing interest in my studies and asking about my progress. Thank you for being supportive of our move to Massachusetts even when you would rather have had us nearer to home. Thank you for coming to visit us and for giving us the pleasure of hosting you and being your tour guides. Thank you for putting up with me during the many times I brought my work along with me on visits.

Most of all, I would like to thank Pam, Lyndon, and Kira. This dissertation and degree would not have happened, Pam, without the many sacrifices you made along the way. I know this is always said in acknowledgement contexts like this. I don't know how often it has been true when others have said it, but it is certainly true in my case. You agreed to move halfway across the continent when you knew how much you would miss the nearness of friends and family. You endured many days and evenings of isolation in our Northampton apartment, looking after Lyndon and Kira while I was working at the office. You switched countless day shifts to night shifts to accommodate my schedule and teaching responsibilities. You tirelessly kept our household in order while I focused on other things. I will be forever grateful. Thank

you Lyndon and Kira for putting up with a dad who spent so many evenings and weekends upstairs in his office staring at his computer. I love you all so much.

ABSTRACT

THE STRUCTURE OF CONSCIOUSNESS

SEPTEMBER 2013

LOWELL KEITH FRIESEN

B.F.A., PROVIDENCE COLLEGE

B.Mus., UNIVERSITY OF MANITOBA

M.A., UNIVERSITY OF MANITOBA

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Joseph Levine

In this dissertation, I examine the nature and structure of consciousness. Conscious experience is often said to be phenomenally unified, and subjects of consciousness are often self-conscious. I ask whether these features necessarily accompany conscious experience. Is it necessarily the case, for instance, that all of a conscious subject's experiences at a time are phenomenally unified? And is it necessarily the case that subjects of consciousness are self-conscious whenever they are conscious? I argue that the answer to the former is affirmative and the latter negative.

In the first chapter, I set the stage by distinguishing phenomenal unity from other species of conscious unity. A pair of conscious states is phenomenally unified if they are experienced together as part of a single experience that encompasses them both. In this and the next two chapters I defend the thesis that, necessarily, for any subject

(of conscious mental states) at any time, all of that subject's conscious mental states (at that time) are part of a single, maximal state of consciousness. I call this thesis the "Unity Thesis." I proceed by considering some preliminary questions that might be raised about the Unity Thesis. For instance, the thesis presupposes that it is coherent to talk about parts of mental states. I consider objections by Tye and Searle and argue that the notion of an experiential part is unproblematic. In the remaining pages of the chapter, I present the source of the biggest challenge to the Unity Thesis: the data gathered from split-brain subjects.

The Unity Thesis is formulated using the notion of a maximal state of consciousness. In the second chapter, I attempt to precisify this notion in a way that does not pre-emptively decide the debate over the Unity Thesis. In informal terms, a maximal state of consciousness is a sum of conscious states that are i) simultaneous, ii) have the same subject, and iii) all have a conjoint phenomenology. I call this the Consensus View. I then consider two unorthodox views that the Consensus View does not take off the table: the views that a "collective consciousness" and a "spread consciousness" are possible. A collective subject is one that can enjoy the experiences of an indeterminate number of "lesser" subjects of consciousness by sharing them together with those subjects. A spread subject is one that can enjoy the experiences of an indeterminate number of lesser subjects of consciousness, but it does so, not by sharing those experiences with the lesser subjects, but by absorbing the lesser subjects of experience into itself, thereby erasing the traditional boundaries between the entities we intuitively think of as subjects of experience. I argue that, although the Consensus View does not decide against them, these views stretch the bounds of coherence and should not, therefore, be accepted. Having presented an account of what maximal state of consciousness is, I define a stream of consciousness in terms of a maximal states of consciousness. In the rest of chapter two, I consider and argue against a number of different ways of interpreting the split-brain data that are either

inconsistent with the Unity Thesis or attribute more than one subject of consciousness to split-brain subjects. Among the views I consider are Lockwood's partial-unity view and the views, by theorists such as Sperry, Koch, Puccetti, Marks, and Tye, that split-brain subjects have two non-overlapping streams of consciousness.

In chapter three, I consider a recent attempt by Bayne to account for the split-brain data in a way that does not attribute two streams of consciousness to them. According to Bayne's Switch Model, the consciousness of split-brain subjects can be likened to that of a ball that is passed back and forth between the two hemispheres of the upper-brain. The hemispheres take turns supporting a single stream of consciousness. I consider the empirical data in some detail and argue that the data is not as compatible with the Switch Model as Bayne claims. I close the chapter by presenting the rough outline of an interpretation of the split-brain data that is consistent with both the Unity Thesis and the split-brain data.

In chapter four, I turn from defending the Unity Thesis to examining an attempt to account for conscious unity. Rosenthal has offered a theory of conscious unity as an extension of his higher-order theory of consciousness. I consider his account of conscious unity in light of a well-known objection to his theory: the (Representational) Mismatch Objection. It can be asked what it is like for a subject of experience when a higher-order state misrepresents its target first-order state. If what it is like for the subject corresponds to the content of the higher-order state, then it appears as though higher-order representation is unnecessary for conscious experience, for it would appear as though it is possible for a state to be conscious without being represented by a higher-order state. If what it is like corresponds to the content of the lower-order state, then it would again seem as though representation at the higher-order level is unnecessary for conscious experience, for the higher-order state would not seem to be doing any work in generating the experience. I consider and argue against two recent defences of Rosenthal's higher-order theory from the Mismatch

Objection. Then I turn to Rosenthal's account of conscious unity. Rosenthal's account posits two mental mechanisms. I refer to the ways of accounting for conscious unity via these two mechanisms as the "gathering strategy" and the "common-ascription strategy" respectively. Both of these strategies, I argue, appear to locate the basis for certain phenomenal facts in higher-order representational facts. This raises a *prima facie* question: does Rosenthal's account of conscious unity land him square within the sights of the Mismatch Objection? Although the gathering strategy may ultimately be understood in a way that does not make it subject to the Mismatch Objection, Rosenthal has certain commitments that bar this strategy from serving as a complete account of conscious unity. This is problematic for Rosenthal, I argue, because his common-ascription strategy faces some difficult questions. This strategy makes conscious unity due to an implicit expectation a subject of consciousness has that, whenever he or she engages in introspection, an explicit sense of conscious unity will be generated. I argue that it is very difficult to see how such an implicit sense could both avoid the Mismatch Objection and do the work it needs to do in order to account for conscious unity.

In chapter five, the discussion turns from the unity of consciousness to self-consciousness. The question that is considered in this and the last chapter is the question whether conscious experience is necessarily accompanied by self-consciousness. The affirmative answer to this question I call the Ubiquity Thesis. I spend some time distinguishing robust conceptions of self-consciousness from minimal conceptions of self-consciousness. The notion of self-consciousness invoked by the Ubiquity Thesis is a minimal one. In spite of the fact that the Ubiquity Thesis invokes only a minimal or thin conception of self-consciousness, I believe the thesis to be false and argue against it. In this chapter I take up the views of Husserl. Husserl is often regarded as the progenitor of the phenomenological tradition, a tradition in which many philosophers affirm the Ubiquity Thesis. I examine and argue against an interpretation of

Husserl's work, one defended by Zahavi, according to which Husserl could be seen to defend the Ubiquity Thesis. One claim that Husserl makes is that, in order for an object to become the intentional target of a conscious state, it must be given to consciousness beforehand. It is possible, during acts of deliberate introspection, for consciousness to take itself as its object. On Husserl's view, this requires consciousness to be given to itself beforehand. This self-givenness of consciousness, argues Zahavi, can be seen as a kind of minimal self-consciousness. Husserl has also offered an account of this self-givenness of consciousness and it appears in his discussion of inner time-consciousness. I attempt to argue, using some of Husserl's other views regarding psychological stances (or standpoints), that consciousness is not given to itself outside of the adoption of a certain psychological standpoint. I also offer an alternative way of accounting for inner time-consciousness, one that does not have, as a built-in feature, that consciousness always has itself as a secondary object.

In the sixth and final chapter, I take up a contemporary defence of the Ubiquity Thesis. Kriegel, a higher-order theorist like Rosenthal, has argued that every conscious state is conscious in virtue of the fact that it represents itself. This self-representation is understood as a kind of self-consciousness and, thus, his theory can be seen as affirming the Ubiquity Thesis. In the first part of the chapter, I take issue with the way in which Kriegel lays out the conceptual terrain. In particular, Kriegel countenances a property he calls "intransitive state self-consciousness." I argue that this way of speaking is confused. I then turn to considering Kriegel's account. Kriegel identifies the species of self-consciousness that pervades all of conscious experience with a peripheral awareness of one's own mental states. I argue that such a peripheral inner awareness does not accompany all of our mental states and, thus, that Kriegel's views do not give us reason to accept the Ubiquity Thesis.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	v
ABSTRACT	viii
LIST OF TABLES	xvi
LIST OF FIGURES	xvii
CHAPTER	
1 INTRODUCTION	1
2 PHENOMENAL UNITY AND SPLIT BRAINS	11
2.1 Varieties of Unity	11
2.2 The Unity Thesis	16
2.2.1 Begging the Question	16
2.2.2 The Thesis	19
2.2.3 Experiential Parts	23
2.3 The Challenge to the Unity Thesis	30
2.3.1 Split Brains	31
2.3.2 A Range of Views	38
3 TWO STREAMS OF CONSCIOUSNESS	41
3.1 Maximal States of Consciousness	41
3.1.1 Subjects, Conjoint Phenomenology, and Maximal States	43
3.1.2 Collective Consciousness	50
3.1.3 Spread Consciousness	55
3.1.4 Streams of Consciousness	58
3.2 Split Brains and Two Streams of Consciousness	61

3.2.1	Partial Unity	62
3.2.1.1	Projecting Oneself into a Partially-Unified Consciousness	65
3.2.2	Two (Disjoint) Streams	70
3.2.2.1	Two (Disjoint) Streams and One Subject	73
3.2.2.2	Being Led Astray by Spatial Unity	76
3.2.2.3	Subjects and Conjoint Phenomenology	83
3.2.2.4	Spread Subjects Revisited	87
3.2.2.5	Two (Disjoint) Streams and Two Subjects	88
4	ONE STREAM OF CONSCIOUSNESS	93
4.1	Bayne's Switch Model	94
4.2	Evaluating the Switch Model	96
4.2.1	Radical Shifts	96
4.2.2	Chimera Studies	103
4.2.3	Simultaneous Dual Voluntary Responses	110
4.3	An Alternative to the Switch Model	117
4.4	Before and After the Experimental Conditions	124
5	HOTS, REPRESENTATIONAL MISMATCHES, AND THE UNITY OF CONSCIOUSNESS	128
5.1	The Mismatch Objection	129
5.2	Responses to the Mismatch Objection	132
5.2.1	Direct Defences	133
5.2.1.1	Rosenthal's Response	133
5.2.1.2	Weisberg's Response	137
5.2.1.3	Representational Mismatch as Partial Representation	140
5.2.2	A Fallback Defence	144
5.3	Rosenthal's Theory of the Unity of Consciousness and the Mismatch Objection	150
5.3.1	Rosenthal's Theory of Conscious Unity	151
5.3.2	The Gathering Strategy and Representational Mismatches	155

5.3.3	The Common-Ascription Strategy and Representational Mismatches	161
6	HUSSERL ON PRE-REFLECTIVE SELF-CONSCIOUSNESS	166
6.1	Self-Consciousness	167
6.2	Husserl and the Ubiquity Thesis	169
6.2.1	Zahavi's Interpretation of Husserl	171
6.2.2	Husserl's Account of Inner Time Consciousness	174
6.3	Objections to Husserl	180
6.3.1	A General Worry	180
6.3.2	Double Intentionality	185
6.3.2.1	The Self-Givenness of Consciousness	190
6.3.2.2	Retaining Consciousness	197
7	KRIEGEL ON INTRANSITIVE SELF-CONSCIOUSNESS	202
7.1	Kriegel's Account of Subjective Character	203
7.2	Objections to Kriegel's Account	207
7.2.1	Is Intransitive Self-Consciousness Really Intransitive?	208
7.2.2	Is Peripheral Inner Awareness or Minimal Self-Consciousness Phenomenologically Manifest?	217
7.2.3	Are We <i>Always</i> in a State of Peripheral Inner Awareness?	220
8	CONCLUSION	229
	BIBLIOGRAPHY	233

LIST OF TABLES

Table	Page
6.1 Consciousness Over Time	177
6.2 Consciousness Without Double Intentionality	199

LIST OF FIGURES

Figure	Page
3.1 Two Overlapping Streams	77
3.2 Two Non-Overlapping Streams	77
3.3 One Stream	78

CHAPTER 1

INTRODUCTION

This is an essay on the nature and structure of conscious experience. One thing that separates us from the many other species we share our biosphere with is our capacity for conscious experience and the comparative richness of our experience. Plants are alive, but they do not have nervous systems. Insects and invertebrates have nervous systems, but in most cases they are vastly more primitive than our own; if they have experiences at all, there is no doubt that their experiential lives are much more impoverished than ours. The nervous systems of some non-human mammals, on the other hand, are comparatively much more complex. But again, there is little doubt that only a small percentage of the non-human mammals on our planet enjoy experiential lives that rival ours in terms of richness and complexity. The comparative richness and complexity of our inner, conscious lives is one of the things that, colloquially speaking, makes us human. This makes the study of the nature and structure of our conscious experience a particularly significant one. An exhaustive study of the nature and structure of conscious experience would take volumes. My project is a relatively modest one. I will explore but two aspects of conscious experience. I will begin by considering the extent to which consciousness is unified and I will end by considering the relationship between consciousness and self-consciousness.

There are a number of distinguishable ways in which our conscious experience seems to be unified. My focus will be on a particular species of conscious unity called *phenomenal unity*. When we introspect, it seems as though all of our conscious states

at a given time form a package of conscious states that are all experienced together at once. To say that a subject's conscious experience is phenomenally unified is, roughly, to say that all of the experiences the subject enjoys at a time belong to a single maximal experience, that there is no experience that the subject has at a time that is not experienced together with all of the other experiences the subject has at that time.

One of the things that makes the study of conscious unity, and, in particular, the study of phenomenal unity, so interesting is that it allows one to grapple with some of the most important and venerable questions in the philosophy of mind. For instance, in theorizing about the unity of consciousness one cannot help but consider questions about how different methods of enquiry and the respective data they deliver are to be reconciled with each other.

Up until quite recently, the dominant method of enquiry employed by theorists involved little more than careful introspection. All that was required to understand the mind, it was thought, was careful scrutiny of one's own experience. Then, setting aside sceptical worries about the existence of other minds, one generalized the observations gathered from introspection of one's own experience to that of others. In recent decades, introspection has seen some of its dominance give way to considerations of empirical data gathered by psychologists and neuroscientists in their laboratories; empirical research is now seen, in many circles, to be at least as relevant to theorizing about the nature and structure of conscious experience as the results of introspection are. In some corners the pendulum has swung even further. Theorists in these circles are inclined to ignore the first-person perspective altogether and, in some cases, even go so far as to disparage it. Evidence from empirical research, together with its apparent implications, is taken at face-value even if doing so involves rejecting what seems to be true about conscious experience based on armchair reflection.

These different approaches to evidence-gathering and the weighting of that evidence are significant. One of the main questions I will take up in this essay has to do with the respect to which the verdicts of armchair reflection on experience are, or, at least, appear to be, at significant odds with the conclusions reached by some psychologists and neuroscientists on the basis of empirical research. For instance, introspection and armchair reflection on our own experience seems to support the view that experience is phenomenally unified. The deliverances of the empirical sciences, however, have led some to espouse and defend views that are incompatible with the *prima facie* deliverances of armchair reflection. These views are views according to which consciousness can fail to be (and often is not) phenomenally unified.

The body of empirical research that has played the most significant role in leading theorists to question the phenomenal unity of consciousness comes from the study of split-brain subjects. Split-brain subjects are individuals that have had their corpus callosum surgically severed. Placed in certain experimental conditions, these subjects behave and react in ways that suggest a lack (or lesser degree) of mental cohesion and unity. Advances in neuroscience in the last few decades are painting a progressively clearer picture of paths that sensory information takes as the brain processes it. What this has shown us is that the two hemispheres of the upper brain are capable of working quite independently of each other. Combining this with the data from split-brain subjects has suggested to some that split-brain subjects may have partially overlapping streams of consciousness or that they have two completely separate streams of consciousness. These views contradict, in a relatively stark way, what we might naïvely believe to be true about the nature and structure of conscious experience.

The situation is one consisting of competing views that are associated, at least to some degree, with different methods of enquiry. On the one hand, we have arm-chair reflection on one's own experience; on the other, we have theorizing on the basis

of discoveries made in laboratories. In the former, the focus is on what it is like from the inside; in the latter, the primary focus is on how the brain works. This is, admittedly, a relatively crude distinction. Few theorists engaging in arm-chair reflection on conscious experience believe even for a minute that conscious experience is not intimately related to and dependent upon what happens in the brain. Likewise, many psychologists and neuroscientists take their discoveries to have either direct or indirect implications for the character and structure of conscious experience. Furthermore, these crudely delineated types of enquiry can be, and, in many cases, already are, blended.¹ Psychologists, on the other hand, do not restrict themselves to merely observing how subjects perform on one task or another. They often explicitly ask their experimental subjects about the nature of their experience. The point, here, is just that, if our attention is restricted to either the deliverances of introspection or the deliverances of empirical research (narrowly construed), we seem to be led to incompatible views about the unity of consciousness and, in particular, the ubiquity of phenomenal unity. Thus, the views one will ultimately be inclined to accept regarding different questions concerning the unity of consciousness will depend, to some extent, on methods of enquiry one accepts as valid and, as a result, the bodies of evidence that one admits as relevant to the questions one takes up.

Only a small minority of theorists nowadays sit at either end of the spectrum, insisting either that only the deliverances of introspection or that only the deliverances of empirical research are relevant to deciding questions about the nature and structure of conscious experience. Most theorists are much more liberal in their approach to evidence-gathering. As has just been pointed out, however, merely adopting a more liberal attitude toward the evidence one admits for consideration does not make the task easier. These theorists must either find a way to show that the conflict between

¹See, for instance, the method of enquiry Dennett (1991) calls “heterophenomenology.”

the data-sets is merely apparent and not a deep one or they must come up with a principled reason to take one body of evidence as more authoritative. Neither of these are easy tasks.

I count myself among those who accept both the evidence from introspection and the evidence from the laboratory. This means that there are decisions to make when the evidence from introspection appears to conflict with evidence from the empirical sciences. I do not have an account of how the two bodies of evidence are to be weighted relative to each other. Nor will I develop one here. My own inclination is to look for a way to reconcile the data in a way that does not require one to declare that one sort of evidence is more authoritative than the other. It would, perhaps, be naïve to enter a study such as this with the belief that reconciliation will always be possible. There may well be situations in which the only way to move forward will be to discount or ignore some evidence in favour of other evidence, and in such situations, it would be good to have a reasoned policy that one could apply consistently. Because I do not have such a policy, such situations, should they arise, will be decided on a case-by-case basis. My goal, however, is to present and defend views that are able to accommodate as much of the evidence as possible.

I admit at the outset that the view regarding the unity of consciousness I defend in this essay is traditional in character and strongly influenced by arm-chair considerations and introspection. This is not because I believe that the deliverances of introspection *trump* the deliverances of the empirical sciences or that data collected in the laboratory are not relevant to the questions I will be taking up. On the contrary, empirical data must be accounted for in our theorizing just as the data from introspection must, and the findings of psychologists and neuroscientists have as much right to shape our theorizing as our armchair reflection on our own experience does. The reasons for the traditional character of the view I defend are twofold. The first is that, contrary to the theorists who have defended non-traditional views

about phenomenal unity, I believe that the empirical evidence can be understood in a way that is consistent with traditional views. I believe the scientific evidence can be taken seriously and accommodated within a view that is traditional in character. The *prima facie* implications of empirical research are not always correct, and, in many cases, the data gained from empirical research are capable of supporting different interpretations. In broad summary, the conclusions I reach in this essay involve, in part, interpreting the empirical data in a way that is consistent with conclusions one might reach by taking introspection-based evidence as authoritative at the outset. This is not to say that the views I defend are merely the result of a desire to preserve traditional views on the subject. For what it is worth, I entered this enquiry open to being convinced that the traditional views on the subject were incorrect. I believe, and will argue, that the interpretations of the empirical data offered by some are problematic in and of themselves. If I am correct, we are left with a situation in which there is an interpretation of the empirical data that is consistent with the view that consciousness is phenomenally unified and that the views according to which consciousness is not phenomenally unified are independently problematic.

The second is that the nature of the questions I take up skew things in a certain way. Consider the general topic of this essay: the nature and structure of conscious experience. The target phenomenon is the inner life, the conscious experiences of conscious subjects such as ourselves. What is under discussion is *what it is like* to be a subject of consciousness. The primary access we have to the phenomenon under discussion is via introspection and reflection. The notion of phenomenal unity, for instance, cannot even be understood unless one is able to pick out the relevant feature in one's own experience. For instance, any discussion of maximal experiences or of experiences being experienced together or, for that matter, of experience itself is meaningless unless all parties have an intuitive grasp of what the referents of the relevant expressions are. And it is doubtful that an intuitive grasp of these notions is

possible without introspecting on the nature and structure of one's experience. Thus, the very topic under discussion is one that requires us to engage in introspection and reflection.

Given that the table is set in this way, one might wonder how the above claim about being liberal with respect to considering evidence from both arm-chair reflection and the empirical sciences and the claim that evidence from the sciences has as much right to shape theorizing as arm-chair reflection will be honoured. There are several ways in which the empirical sciences have effected change in the philosophy of mind in the last two or three decades. Perhaps one of the most significant changes is that philosophers of mind no longer place as much confidence in the deliverances of introspection as philosophers of mind used to do. The reason for this is that empirical research on the mind has shown us that we are often not very reliable—or, at very least, not nearly as reliable as we believe ourselves to be—when it comes to reporting on our own internal states. This means that the deliverances of introspection do not get a free pass. They must be scrutinized carefully, and then we must ask ourselves whether our *prima facie* judgements about our own minds might be incorrect. What we have learned is that there is no guarantee that our subjective sense of our own internal states always delivers accurately. Thus, even though the nature of the questions I will be taking up in this essay afford a certain status to introspection and reflection, there is still need to take empirical research seriously and to carefully consider the implications of that research.

I began this introduction by saying that one of the things that makes a study of the nature and structure of conscious experience so interesting is that it affords one the opportunity to grapple with some of the most important and venerable questions in the philosophy of mind. Given how much we have learned and how much we continue to learn about how our subjective sense of our own internal states can fail us, the question regarding how to reconcile these different methods of enquiry and

their associated bodies of evidence when they appear to conflict is one of the most important questions in the philosophy of mind. As I have suggested above, I do not have a general answer to this question and I will not be working out such an answer in this essay. I will, however, be grappling with this issue directly by attempting to reconcile seemingly conflicting bodies of evidence. The apparent conflicts between these bodies of evidence are also the source of interest and intrigue. The fact that the deliveries of introspection and the deliveries of empirical research do not always mesh in a straightforward and obvious way is a significant and interesting discovery in itself, but learning about the many different ways in which the empirical research seems to conflict with our naïve beliefs about our own minds continues to reward the curious and anyone who seeks to understand.

Another question that an enquiry into the nature and structure of conscious experience affords one the opportunity to grapple with has to do with the relationship between the mental and the physical. Perhaps the first question that comes to mind is the question associated with the so-called hard problem of consciousness (Chalmers, 1995): why do brain states give rise to experiential states? Another question having to do with the relationship between the mental and the physical has to do with understanding the way in which activity in different parts of the brain affects conscious experience. These are independent questions. Whereas the empirical sciences may someday provide definitive answers to the latter, it is an open question whether neuroscience, even completed neuroscience, will ever be able to answer the former (Levine, 2001). Although an enquiry into the nature and structure of conscious experience may well influence one's stance on the hard problem of consciousness, the question of the relationship between the mental and the physical that it engages more directly is the question regarding the roles different parts of the brain have with respect to determining the structure and character of conscious experience. For instance, some of the non-traditional views regarding phenomenal unity I will be arguing against in

this essay are motivated, not just by the behavioural responses of subjects in certain experimental conditions, but by what we have learned about brain architecture and functional organization. The fact that certain parts of the brain seem capable of functioning and processing information independently of other parts is taken as support for various claims to the effect that brains are capable of housing more than one mind or stream of consciousness. It would be a significant result if it were determined that a single brain is capable of giving rise to multiple minds or streams of consciousness.

There are a number of questions one might have about phenomenal unity. How prevalent is it? Is phenomenal unity merely a property of the experiential lives of normal human subjects or is consciousness also phenomenally unified in certain pathological cases and non-normal subjects? Is consciousness necessarily phenomenally unified? These questions will be addressed in the first three chapters. Having argued that consciousness is phenomenally unified, I will turn my attention to a related question: if consciousness is unified, how is this property of consciousness to be accounted for? Like many of the other questions I take up in this essay, this question could take up an entire study in itself. My exploration of this question will be restricted to evaluating one answer to it. Rosenthal has theorized about many aspects of consciousness and is, perhaps, most well-known for his higher-order theory of consciousness.² He has recently attempted to account for conscious unity within his general theory of consciousness. In chapter 5, I will examine his views and I will argue that some of the mechanisms he invokes to account for conscious unity expose his higher-order theory of consciousness to a well-known objection.

In the last two chapters of this essay, the focus will turn from the unity of consciousness to self-consciousness. Many theorists, particularly those working within the phenomenological tradition, argue that it is impossible for a subject to be conscious

²See his 1986 and his 1997.

and fail to be self-conscious. It is natural to discuss the views of these theorists in this context, because self-consciousness is an inherently reflexive phenomenon; to be self-conscious is to be aware of oneself in one way or another. Higher-order theories of consciousness, like the one defended by Rosenthal, make conscious experience dependent upon the representational relationships between mental states. These theorists maintain that conscious experience is the result of our awareness of our own mental states. In this regard, the views of higher-order theorists and theorists working in the phenomenological tradition have a lot in common. The phenomenologists who maintain that consciousness is necessarily accompanied by self-consciousness are of this opinion, in part, because they understand consciousness and self-consciousness to consist of our awareness of our own mental states. A mental state is conscious, these theorists say, in virtue of its subject being aware of it. And if a subject is aware of its mental states in this way, it can be said to be self-conscious. In the last two chapters, I will take up the views of two theorists who maintain that conscious experience entails self-consciousness. Husserl is often considered to be the first phenomenologist, and so a study of his views is important for understanding the views of the many other phenomenologists of the 20th century. His views will be examined in chapter 6. In chapter 7, I will take up the views of Kriegel. Kriegel has defended a view that might be considered a contemporary version of Brentano's theory of consciousness. Brentano was the philosophical mentor of Husserl. Kriegel's theory of consciousness is also a higher-order theory of consciousness and so has much in common with Rosenthal's theory. The conclusion I will argue for in the last two chapters is that Husserl and Kriegel have not given us good reasons to believe that consciousness entails self-consciousness.

CHAPTER 2

PHENOMENAL UNITY AND SPLIT BRAINS

Philosophers have distinguished several different varieties of conscious unity. The question I will seek to answer in this and the next two chapters is whether or not, and to what extent, conscious experience is *phenomenally* unified. I will have more to say about this question momentarily, but before doing so, I will introduce the notion of phenomenal unity by contrasting it with some other species of conscious unity that theorists have identified.

2.1 Varieties of Unity

One sort of unity exhibited by our experience involves the way in which we perceive objects. As I look at the coffee mug sitting on my desk, I perceive redness and cylindricity. If I hold my hand near it, I experience warmth. And if I grab its handle and pick it up, I feel weight and a slight resistance against the direction I wish to move my arm. What is noteworthy about my experience is that my experience is of a single object in which several properties inhere. I experience its colour, its complex shape, its warmth and weight as all inhering in the same object, my mug. I do not, for instance, experience redness as belonging to one object and cylindricity as belonging to another, even though it is possible to have an experience in which this is the case—imagine gazing at a red cube beside a blue cylinder. Why do we sometimes experience redness and cylindricity as inhering in the same object and not at others? This question is all the more pressing when we consider the fact that empirical research into vision has shown that different properties of objects are

processed by different parts of our visual system and in different regions in the brain. Attempts to provide an answer to this question have come to be known as solutions to the *binding problem*.¹

The binding problem, at least as it is most commonly discussed, is a problem having to do with the visual system and visual perception. But vision is not the only sensory modality for which a kind of binding problem arises. Consider the sense of hearing. When listening to a piece of music, one can isolate and attend to the different sounds made by the different instruments; one can, for instance, focus on the vocalist or on one of the accompanying instruments. Yet the sounds made by the different instruments and by the vocalist are complex sounds. They each have a pitch and they each have a particular timbre and character that we associate with the objects that produce them; the sound produced by the vocalist, for instance, has a vocal character and the sound produced by the piano has a piano character. This is what makes them different sounds, and they are distinguishable even if they happen to be sounds of the same pitch, as they are, for instance, when the voice and the piano happen to carry the melody simultaneously. When listening to a piece of music, some properties are experienced as belonging to the same sound, and some are experienced as inhering in different sounds. So just as there is a binding problem with respect to visual perception, there is a binding problem for aural perception.²

Both these binding problems are *intramodal*. However, as my perception of my coffee mug demonstrates, binding problems can also arise *intermodally*. Our experience of the world around us is rich and complex. We experience the objects in it

¹For a concise introduction to the binding problem, see Roskies (1999). Much of the pioneering work on this problem has been done by Treisman. For a concise overview of some proposed solutions to the problem, see her 1999.

²The visual and hearing systems are different systems and so it is quite possible, indeed, likely, that auditory stimuli are processed differently than visual stimuli. Thus, the neurological considerations that come to play with respect to visual binding may be different than the ones that come to play with respect to auditory binding. The binding problem for vision may be quite different than the binding problems that may arise in other sensory modalities.

as possessing many different properties, and often the properties we experience them as having are properties we perceive via different sensory modalities. This kind of experiential unity is known as *object unity*.³ Elements or parts of an experience are object-unified when they are experienced as properties of a single object. Experiences are *phenomenally unified* when they are experienced together as part of a larger encompassing experience. Consider again my experience of my coffee mug. While I look at the mug, I also hear music playing. Both my experience of the mug and my experience of the music are part of a single global experience. If I happen to be aware of the feeling of the chair on my back at the same time as I am aware of the mug and the music, then the feeling of the chair is also part of the same global experience.

What does it mean for a number of experiences to be part of a single global experience? Many theorists who consider the notion of phenomenal unity trust that the idea of a number of experiences being part of the same global experience is intuitively clear. Others attempt to explicate the notion by saying that experiences are phenomenally unified when they have a *conjoint phenomenology*. Philosophers often characterize the difference between conscious and unconscious mental states by pointing out that, for conscious mental states, there is “something it is like” for a subject to be in them (Nagel, 1974). They have a “raw feel.” Unconscious mental states do not have this feature. There is something it is like for me when I look at my mug. Likewise, there is something it is like for me to hear the music playing. These are both conscious states, and both of them can be had independently of each other. But when I see my mug and hear the music at the same time, the what-it’s-like-ness of the two experiences come together into a single experience whose phenomenal character contains both the what-it’s-like-ness of seeing the mug and the what-it’s-like-ness of hearing the music. Thus, there is something it is like to see the mug, there

³See Tye (2003). Bayne and Chalmers (2003) call this kind of unity “objectual unity.”

is something it is like to hear the music, and there is something it is like to see the mug and hear the music at the same time. The latter is distinct from each of the former. When a pair of experiences come together and each contribute to the overall what-it's-like-ness of a single global state, they have a conjoint phenomenology. Theorists have introduced various technical terms for the relation that obtains between a pair of experiences when they are each part of the same experience in this way. Parfit (1984) and Lockwood (1989) say that such experiences are *co-conscious*. Bayne and Chalmers (2003) and Bayne (2008, 2010) say that such experiences are *subsumed* by the experience of which they are both a part. I have been following Tye (2003) in saying they are *phenomenally unified*.

Both phenomenal unity and object-unity are species of conscious unity that experiences may have at a time, i.e., synchronically.⁴ Conscious experience, however, may also be unified diachronically. My current experience, for instance, is experienced as part of temporally extended experience. It is part of an experiential stream or stream of consciousness. When experiences are unified in this way, they are *stream-unified*. In addition, beings like us, who have the capacity to ascribe properties to objects and individuals, can ascribe ownership to experiences. I can, for instance, think of the experience I have as I reach for my mug as my own. And not only can I self-ascribe my present experiences, but I can also do so for experiences that I have had in the past. When experiences are unified in this way, they are *subject-unified*.

One might wonder whether experiences could ever be stream-unified and fail to be subject-unified. The answer to this question hinges upon the way in which one conceives of a subject of experience and upon the way in which the notion of subject-

⁴Most cases of object unity are cases of synchronic unity, but there may be some instances in which experience is object-unified over time. Consider, again, the case of listening to music. Melodies and musical sounds are temporally extended objects, and we experience their different properties as inhering in them over time. Thus, it seems as though sometimes object unity is not merely a sort of synchronic unity.

unity is specified. As it is presented above, subject-unity involves a relatively robust notion of a subject, one according to which a subject of experience must be capable of self-ascribing mental states. The capacity to self-ascribe mental states requires a significant degree of cognitive sophistication. It is quite possible that we humans are the only inhabitants of our planet that have this capacity and we attain this capacity only when we reach a certain stage in our cognitive development. If a creature has experiences but does not have the capacity to self-ascribe mental states, then the experiences of such a creature could be stream-unified without being subject-unified.⁵ A subject suffering from multiple-personality disorder might constitute another case in which stream- and subject-unity come apart. Such an individual might fail to ascribe a pair of experiences to the same subject even though the experiences both occur within the same stream. Multiple-personality disorder is a complex phenomenon and one that is subject to a variety of competing interpretations. Whether or not individuals suffering from multiple-personality disorder constitute examples of cases in which stream-unity and subject-unity come apart is a question I do not have the space to address fully here. A lot would ride on one's account of how mental states refer to objects.

But there are other ways of conceiving of subject-unity. Some have suggested that the having of an experience has its own distinctive phenomenal feel. On this view, not only does my experience of my mug have a reddish phenomenal feel, its total phenomenal character includes an additional feel of being mine. One might call this the feel of ownership. If the notion of subject-unity is specified in a way that makes the feel of ownership but not self-ascription a requirement on subject-unity,

⁵Many philosophers accept that this possibility is also an actuality. There are some, however, who would deny this. According to some philosophers (Carruthers, 1996, 2000; Dennett, 1991), the having of experiences is dependent upon having certain linguistic, conceptual, and cognitive capacities. Since the self-ascription of mental states would seem to require these capacities, these philosophers would deny that this possibility is an actuality.

then the question whether it is possible for experience to be stream-unified and not subject-unified would turn on the question whether the feel of ownership necessarily accompanies experience. On the other hand, if all that is required for experiences to be subject-unified is that they belong to the same subject whether or not there is an accompanying feel of ownership, then it is difficult to see how experiences could be stream-unified and fail to be subject-unified. There are deep conceptual connections between the notion of a stream of consciousness and the notion of a subject, and it is quite plausible that a necessary condition on a pair of experiences belonging to the same experiential stream is that they have the same subject.

The notion of a subject of experience will be discussed in more detail in chapter 3. What will be made clear there is that the notion of a subject of experience that I will be working with in this essay is a relatively bare or thin one. I will not, however, have much more to say about the relationship between stream- and subject-unity.

2.2 The Unity Thesis

In the introduction and in the beginning of this chapter, I said that the question I would be considering here is the question whether or not, and to what extent, consciousness is phenomenally unified. Having distinguished phenomenal unity from other sorts of conscious unity, I will now attempt to make clear exactly what is being asked. Before doing so, however, it will be necessary to address a problem that might be raised in response to the way the discussion has proceeded thus far.

2.2.1 Begging the Question

The problem might be stated in the following way. The project is that of deciding whether or not consciousness is phenomenally unified. The way in which the notion of phenomenal unity was presented presupposes a certain answer to the question. In attempting to communicate what is meant by the notion of phenomenal unity,

the reader was pointed to a feature of his or her own experience and told that, when consciousness has that feature, it is phenomenally unified. Thus, in describing phenomenal unity, the question was already decided; the question is not an open one, because the notion of phenomenal unity was characterized in such a way as to determine a certain answer to the question. My experience of my mug and my experience of the music were phenomenally unified because they were experienced together as part of a global experience at that time. Given that I seem to be able to point to instances of phenomenal unity, it would seem that the answer to the question is affirmative: consciousness is phenomenally unified.

There are a couple of reasons why introducing phenomenal unity in this way is not as problematic as it might seem. First, it is possible to question whether appearance matches reality. There is relatively little doubt among theorists that, for most of us, our conscious experience *seems* to be phenomenally unified, but one might wonder along with Rosenthal (2005c) whether our conscious experiences are *actually* phenomenally unified or whether they merely appear to be. Given that this is an open question, merely pointing to putative instances of phenomenal unity would not be enough to conclude that consciousness is in fact phenomenally unified. Rosenthal seems prepared to treat this possibility as irrelevant, saying that, when it comes to experience, appearance may be all that matters. According to some of the views I will be considering in the following pages, however, appearance is not all that matters. According to these views, it is to be expected that the individuals for whom consciousness is phenomenally disunified will be unaware of this fact. For these subjects it will seem as though their experiences are phenomenally unified when there is more to the situation than is suggested by their subjective sense of their experience. Thus, according to the defenders of these views, the mere fact that it seems as though experience is phenomenally unified is not enough to establish that it in fact is.

Second, the intended scope of the question would need to be settled before it could be determined whether a simple pointing to a putative case of phenomenal unity would settle it. The question whether consciousness is phenomenally unified certainly seems to be somewhat general in its intended scope, but one could have a number of different questions in mind in asking it. For instance, although the question might not be entirely in keeping with its seemingly general nature, one might ask the question restricting the scope merely to one's own experience at a particular time. Then, if it seemed that one's experience was phenomenally unified and appearance was consistent with reality, then recognizing the apparent phenomenal unity of one's own experience would yield an answer to the question. On the other hand, one might intend the question to have broader scope. One might ask the question wishing to learn whether one's experiences were phenomenally unified over time and not just during periods of reflection. The mere fact that one's experiences appear to be phenomenally unified whenever one introspects would not entail that they have the property as a general matter of course. One could generalize the question further and wonder whether consciousness is phenomenally unified for all conscious creatures like us with nervous systems similar to ours. Then, of course, merely pointing to a putative instance of phenomenal unity in one's own experience would not settle the question. As will become clear shortly, there are some real hurdles to generalizing from one's own experience to that of others. Finally, one might intend the question to have even broader scope. One might wonder not just about the inner lives of creatures in the actual world that are physically similar to us, but also about the inner lives of all conscious creatures, whether they are physically similar to us or not and whether they reside in the actual world or in some merely possible world. Again, in this case merely pointing to a putative instance of phenomenal unity in one's own experience would not settle the question. Thus, there is no deep worry about the way in which the phenomenon of phenomenal unity was introduced.

2.2.2 The Thesis

In what follows, I will understand the question in its widest possible scope. I will take it to concern all actual and merely possible subjects of experience. Even when understood in this perfectly general way, I believe the answer to it is affirmative. That is, I take the following thesis to be true:

The Unity Thesis Necessarily, for any subject (of conscious mental states) at any time, all of that subject's conscious mental states (at that time) are part of a single, maximal state of consciousness.

There are several things that need to be said about the Unity Thesis. First, the notion of consciousness invoked in the Unity Thesis is that of *phenomenal* consciousness. Thus, the expression “conscious mental state,” as it appears in the Unity Thesis, is to be taken as a synonym for “experience.” It is absolutely crucial for what follows to distinguish between phenomenal consciousness and what Block (1995) calls *access*-consciousness. A mental state is phenomenally conscious only if it has a phenomenal character, only if there is something it is like for its subject to be in it. A mental state is access-conscious if it is available to certain mental processes so that it is poised to influence its subject's thoughts and actions. To say that a mental state is access-conscious is just to say that it stands in certain computational or functional relations to other mental states and mental processes. There is by no means a consensus on this, but many believe that it is possible for these notions to come apart, that it is possible for a state to be access-conscious without also being phenomenally conscious and vice versa. This is not to say that there is not a strong correlation between the two. This is evidenced, in part, by the fact that it is difficult to produce clear cases of dissociation (in either direction). The tangled relationship between phenomenal and access-consciousness will be on full display in the following chapters. Much of the debate over the Unity Thesis and the split-brain research that is offered as a challenge to it proceeds by making certain inferences about an individual's phenom-

enally conscious states based on what his or her behaviour suggests about his or her access-conscious states. That is, many who use split-brain research to challenge the Unity Thesis mount their challenge by relying, to greater and lesser extents, on the assumption that facts about access-consciousness tell us in a straight-forward way what the facts regarding phenomenal consciousness are. My own proposal about how the split-brain research is to be interpreted involves, in part, a rejection of this assumption; it dissociates phenomenal consciousness from access-conscious in certain ways.

Second, when the notion of phenomenal unity was introduced in the previous section, it was introduced as a property individual experiences have when they are parts of the same experience. But I have also been speaking somewhat loosely about *consciousness* (as a whole) being unified. As I have been using the expression, to say that the consciousness of a given subject is phenomenally unified is just to say that the subject satisfies the Unity Thesis. It is just to say that all of his or her experiences or conscious mental states at a time are mutually phenomenally unified with each other.

This does not mean that the expression “phenomenal unity” is being used in two different ways. The notion of parthood invoked by the Unity Thesis is just the notion of parthood that appears in classical mereology. In classical mereology, an object can stand in the parthood relation to itself. Of course, in ordinary usage, the term “part” is often not used in this way—if it is said that *a* is part of *b*, it is implied that, if *a* were “removed” from *b*, there would be a remainder. In order to accommodate this use of the term, classical mereology distinguishes between proper and improper parts. An object is an improper (but not a proper) part of itself. Since a maximal experience is a part of itself, it is part of a single, maximal state of consciousness and, thus, is phenomenally unified with itself. The expression “phenomenal unity” is not used in a different sense when it is applied to the sum of a subject’s conscious mental

states or consciousness as a whole at a time. A maximal state of consciousness is just the limit case of phenomenal unity.

Third, the notion of a subject, as it is invoked by the Unity Thesis, is a relatively thin notion. Although our paradigm subjects of experience are all housed in physical bodies, it would be problematic to take the Unity Thesis to be a thesis about biological organisms. For then a pair of Siamese twins that could not be surgically separated would count as a counter-example to the Unity Thesis. Presumably, what we want to say about a pair of Siamese twins like this is that there is one biological organism, one body, that houses two subjects of consciousness, not that the body houses only one subject of consciousness.⁶ What this means is that the notion of a subject that is invoked by the Unity Thesis must be one that does not count subjects by counting bodies. What we need is a thinner notion. The notion need not be so thin as to be the notion of a *mere* subject of a experience, for in much of the debate about whether or not conscious experience is always phenomenally unified the subjects of experience are taken to have mental states such as intentions and memories in addition to conscious sensory states. Thus I will stipulate that the notion of a subject invoked by the Unity Thesis is one that allows for these additional mental states. However, the notion I will be working with in the following pages is not one that makes these additional mental states a necessary condition on being a subject of experience. Mere subjects of experience count as subjects of experience for the sake of the Unity Thesis.

The Unity Thesis is an extremely strong claim, and it may not be possible to establish it conclusively. Although we are certainly familiar with a relatively rich variety of organisms that we presume have inner experiential lives of some sort, this is a presumption that is made on the basis of the fact that they have similar nervous systems to ours. Whether there are other creatures elsewhere in our universe that

⁶Depending on what the notion of a body is, there might also be cases in which it is indeterminate whether a pair of Siamese twins constitute one or two bodies.

have experiential lives we can only speculate. If there are such creatures, we have no way of knowing the extent to which their physical constitution is like ours. Our access to other possible worlds in the modal universe is no more robust than our access to conscious life-forms on planets that are merely spatio-temporally removed from us. Thus, the only way the Unity Thesis could be established would be by some transcendental argument. Although some arguments that will be presented in the next chapter have somewhat of a transcendental flavour, they are directed, first and foremost, at discrediting some views that are inconsistent with the Unity Thesis. If the arguments are successful, they will show that some interpretations of split-brain research that are inconsistent with the Unity Thesis are untenable. Even if these arguments are successful, there is much that would have to be argued before one could be so bold as to claim a proof of the Unity Thesis. I believe the arguments I will present in the next two chapters go some distance to making the Unity Thesis more credible than its denial, but I make no claim to having proved the Unity Thesis.

My project in this and the following two chapters is more of a defence of the Unity Thesis from putative challenges to it than it is to make a positive case for it. To my knowledge, the most significant challenges to the Unity Thesis are rooted in the empirical literature on split-brain subjects. Data gathered from observing these subjects in certain controlled situations have been used by theorists to offer and defend various views that are inconsistent with the Unity Thesis. I believe that the split-brain data does not require one to abandon the Unity Thesis. I will defend the Unity Thesis by arguing against these views and by showing that the split-brain data can, in fact, be made consistent with the Unity Thesis. Thus, although the thesis I will be defending makes a claim about what is necessarily true of consciousness, my defence of it will focus, largely, on what seems to be true of the consciousness of human beings (and certain non-human animals) in our relatively small corner of the actual world.

The plan for the rest of this chapter is as follows. In section 2.3, I will present some of the empirical research on split-brain subjects that has led many theorists to doubt that phenomenal unity is as ubiquitous as one might be inclined to believe on the basis of one's experience. Then I will outline some interpretations these theorists have offered of the split-brain data. This will set the stage for chapter 3, in which I will argue against them. Before moving on to the split-brain research, however, I will consider an objection to the Unity Thesis that questions an aspect of the conceptual framework within which the thesis is formulated.

2.2.3 Experiential Parts

According to the Unity Thesis, experiences are phenomenally unified whenever they are part of the same maximal experience or state of consciousness. This way of formulating what it means to say that consciousness is phenomenally unified depends, crucially, on the notion of an experience being *part* of another experience. Several theorists have argued that the notion of an experiential part is problematic. William James (1981, ch. 6) argued this in the late 19th century, and has been followed more recently by Carnap (1967), Searle (2000), and Tye (2003). Many theorists who have taken up the subject of phenomenal unity in recent years have, contra James, Tye, and Searle, taken the notion of an experiential part to be a coherent notion and, thus, have seen fit to include the notion of an experiential part within the conceptual frameworks they employ in their discussions of phenomenal unity. Some of them have offered what are, in my opinion, sound counter-arguments to those of James, Tye, and Searle. I will only rehearse the issue briefly here.

Tye argues that, if the question of phenomenal unity is posed as a question regarding how distinct experiences are unified as parts of a single experience, certain infinite-regress problems lurk. Phenomenal unity makes a difference with respect to phenomenal character; there is a phenomenal difference between experiencing a num-

ber of experiences, $e_1 - e_5$, separately and experiencing them together in a unified experience. Where could this difference in phenomenal character come from unless it came from the fact that the unifying relation, R_1 , between $e_1 - e_5$ has its own phenomenal character and is itself experienced? But the phenomenal character of R_1 is also part of the subject's unified experience at the time. Thus, there must be a second unifying relation, R_2 , that unifies the phenomenal character of R_1 with the respective phenomenal characters of $e_1 - e_5$. But R_2 "must itself be experienced. For the unity is phenomenal. And now a regress has begun to which there is no end" (2003, 22).⁷

Bayne (2010, 30–32) has offered several responses to this argument. The first is that, if the argument is sound, then it would seem to threaten *any* account of phenomenal unity. If whatever is responsible for unifying experiences has its own unique phenomenal character that is itself experienced, then its phenomenal character must be unified with whatever other experiences a subject has at a time. As soon as this is granted, the regress is up and running. Even Tye's own preferred account of phenomenal unity seems subject to this kind of regress problem. Tye argues that phenomenal unity is the result of distinct phenomenal contents entering into a single phenomenal content. Since phenomenal unity makes a phenomenal difference, the question could be asked how this phenomenal difference is accounted for within his account. If the phenomenal difference must come from the experience of a special phenomenal character over and above the phenomenal characters of the other elements in the experience, then this special phenomenal character must be accounted for within the content of that experience. But then this additional content must be unified with the rest of the contents of the experience, and a regress is up and running.⁸

⁷See also Siewert (2001).

⁸Hurley (1998), though not responding to Tye directly, offers similar arguments against content-based accounts of unity.

A second response offered by Bayne is one that rejects the notion that the relation unified experiences stand in when they are unified must have a special phenomenal character that is itself experienced. “Phenomenal unity is a phenomenal relation in the sense that it makes a phenomenal difference, but not in the sense that it has its own phenomenal character that makes an *additional* contribution to what its like to be the subject in question” (original emphasis) (2010, 31). Of course, the challenge for the proponent of this response is to say something about how phenomenal unity can make a phenomenal difference without being committed to saying that the relation responsible for the unity has its own unique phenomenal character. Bayne suggests that the phenomenal difference can be thought of as a *way* of experiencing. One can have some experiences separately, or one can have them together. There is a phenomenal difference between having the experiences separately and having them together, but the phenomenal difference does not have to be a special phenomenal character that is over and above the phenomenal characters of the experiences that are had together.

Searle argues against what he calls the “building block model” of consciousness. This model of consciousness, says Searle, consists of the idea that “any conscious field is made up of its various parts” (2000, 570). Searle’s main objection to this model is with the approach it suggests to the quest for finding the neural correlates of consciousness. If a subject’s conscious field or maximal experience at a time has parts, then it would seem that the way to go about finding the neural correlate(s) of consciousness would be to tackle conscious experience in a piecemeal fashion.⁹ One might begin by finding the neural correlates to visual experience. Then, once the correlates

⁹Roughly, the quest for the neural correlate(s) of consciousness is an effort to find the parts of the brain and nervous system that are most directly responsible for conscious experience. According to Koch (2004), who has made one of the most significant recent contributions to the quest for the neural correlate(s) of consciousness, the goal of the quest is to discover “the minimal set of neuronal events and mechanisms jointly sufficient for a specific conscious percept” (2004, 16). Also, see Chalmers (1996).

of visual experience have been found, tackle the other sensory modalities in the same way. Once the neural correlates of conscious experience in all of the sensory modalities have been discovered, the general quest for the neural correlate(s) of consciousness would be completed; the neural correlate(s) of consciousness as a whole would be the sum of the sense-specific neural correlates. The problem with this piecemeal approach to finding the neural correlate(s) of consciousness, argues Searle, is that it suggests that modality-specific sensory experiences are independent and isolated in ways that they do not seem to be. “The building block theory predicts that in a totally unconscious patient, if the patient meets certain minimal physiological conditions (he is alive, the brain is functioning normally, he has the right temperature, etc), and if you could trigger the (neural correlate of consciousness) for say the experience of red, then the unconscious subject would suddenly have a conscious experience of red and nothing else” (2000, 573). This prediction, argues Searle, is problematic because the picture painted by empirical research suggests that subjects must reach a certain general state of consciousness. Their brains must reach a minimum threshold of general arousal and neuronal activity before any conscious experience is even possible. The switch for (general) consciousness must, so to speak, be turned on before any particular sense-specific conscious experience is possible.

Searle’s argument against the building block model appears to be an argument against a certain autonomy or independence claim: namely, that it is impossible for creatures like us to have certain relatively minimal sense-specific experiences without a more general conscious field being activated. Because visual experience of any kind requires a general level of consciousness, it would seem to be impossible for us to have a visual experience of red and nothing else. The visual experience of red would be accompanied by conscious experience of whatever other sensory information was being received by one’s other sense organs at the time.

There are a number of things to say about Searle's argument. First, the point he makes about whether an otherwise unconscious human subject could be made to have an experience of red and nothing else is well-taken. On this point, even those who Searle counts as proponents of the building block model would probably agree (see Crick and Koch, 1995, 1998; Koch, 2004). In spite of this, and this is my second point, it is not clear that it would be impossible for a human subject to have nothing but an experience of red. Suppose it is sometime in the future and neuroscientists are able to completely isolate and identify the various neural pathways along which sensory information enters the brain. Suppose, furthermore, that these neuroscientists have the capacity to mute the neurons along these pathways without seriously compromising the subject and that the subject can be induced to enter a state of complete mental calm, one in which the subject does not have any conscious thoughts—conscious thoughts, after all, have a phenomenology. It is not at all obvious to me that these neuroscientists would not be able to take a subject like this and, if they knew where to stimulate the brain, produce a conscious experience of red and nothing else.

It must be admitted that the situation is a bit more complicated than this imaginary scenario suggests. One reason for this is that there is some debate about whether general mental arousal has its own distinctive phenomenal feel. If it does, and if the having of a sensory experience requires, as Searle argues, a level of general arousal, then it would be impossible to have an experience of red and nothing else. I am of the view that mere mental arousal does not have a distinctive phenomenal feel. It seems to me that a subject in a general state of mental arousal but with no sensory experience or conscious thoughts would not, at that time, be the subject of experience; it would not be like anything to be that subject. I admit, however, that this claim is tendentious and that there may be no way to decide it conclusively. The main reason for this is that a state of mental arousal apart from any sensory and proprioceptive

experience and apart from any inner dialogue or conscious thought is introspectively inaccessible to us. None of us have ever experienced such a state (or if we have, it occurs for only a very brief period of time as we wake from unconsciousness) and if we were to experience such a state, reflecting upon it and thinking about it would, by the very act, contribute additional phenomenal elements to it. That said, I am sceptical that mere conscious arousal has a distinctive phenomenal feel. It seems to me that a state that is void of any and all of the experiential elements that we typically think of as having a phenomenal feel is not a state for which there would be anything it is like to be in.

If this granted, the question then concerns whether it is possible to have a certain kind of simple experience like an experience of red and nothing else. Even though there is room to doubt its nomological possibility, I believe its metaphysical possibility beyond doubt. Searle may be right about sense-specific experience being dependent upon a certain minimal level of neural activity in the areas of the brain responsible for wakefulness and general consciousness, but, if the argument was intended to support a conclusion about the impossibility of having a certain relatively minimal sense-specific experience like the experience of red by itself, then it seems to me that Searle's argument is far from conclusive.¹⁰

One could point out that the mere fact that Searle has not established the impossibility of the kind of simple experiences discussed above does not entail that they *are* metaphysically possible. Is it true that all experiences enjoy this kind of robust metaphysical independence? It is not clear that they do. Consider the Müller-Lyer illusion. It is arguable that, while looking at a printed image of the two lines, one's experience consists of more than just visual experiences of each of the two lines. The illusion of unequal length is a part of one's experience, but it is not identical to the

¹⁰To be fair to Searle, it is not clear whether Searle intends to establish anything more than a nomological dependence.

visual experience of either of the two lines. Could a subject of experience have the illusion of unequal length independently of the visual experiences of the two lines? It is not clear to me that one could. Or consider what it is like to hear two notes simultaneously. Depending upon their pitches, the two notes can be heard as pleasant, complementary, and harmonious or as discordant. It is not clear, perhaps even doubtful, that the experience of harmony or of discord could be had independently of the experiences of the notes themselves. Although Searle has not established that experiences are dependent upon each other in any robust way, there may be some experiences that cannot be had independently of certain other experiences.

Does this establish that the notion of an experiential part is problematic? I do not think it does. The reason for this is that merely deciding questions concerning whether and to what degree experiences are autonomous and independent of each other is not sufficient to decide questions about whether experiences *have* parts or about whether it is coherent to think of experiences as having parts. This seems to be the upshot of Bayne's criticism of Searle's argument. Bayne (2010, 35–36) responds to Searle by insisting that countenancing experiential parts does not, by itself, commit one to any further claim about the relationship the parts of an experience have to the whole. It is perfectly consistent to maintain that experiences have parts *and* that the parts of some experiences are metaphysically or ontologically dependent upon the wholes of which they are parts or upon other experiences that are part of the same whole.

So far this discussion of experiential parts has consisted in criticizing arguments that aim to show that there is some problem with the notion of an experiential part. I wish to conclude by saying something in support of the notion of an experiential part and by saying something about why this debate is of only secondary importance for what follows. Few, if any, of our experiences are completely homogeneous. Even if experiences can not be sectioned or partitioned in any real way, the fact that they are

not homogeneous makes it possible to carve them up in thought. We can isolate our attention to particular aspects of an experience. This, it seems to me, is enough to say that experiences have parts. Of course, this notion of parthood comes divorced, as Bayne has argued, from various other relations of fundamentality and autonomy, but this relatively bare notion of parthood is a perfectly serviceable notion and has been employed by philosophers in other domains of enquiry. No doubt, there will be some that will be relatively unmoved by this (admittedly thin) argument. In the face of this response, I am content to disagree and move on. The reason for this is that, even if it should turn out that the notion of an experiential part is problematic or that there are no experiential parts, the substance of the following pages will not be significantly effected. Of course, if there are no such things as experiential parts, the Unity Thesis, as I have formulated it, is in trouble, for it is formulated in such a way that it presupposes talk of experiential parts is admissible. But even if the Unity Thesis fails as a means of stating the thought or idea it is intended to capture, this would not entail that the underlying notion that consciousness is phenomenally unified is false. If there are no such things as experiential parts, one would have to find a different way of formulating the thesis, but one would not, thereby, have to give up on the notion of phenomenal unity. As evidence of this, consider the fact that Tye, who denies experiential parts, does not deny that consciousness is phenomenally unified (at least in normal subjects). One might also have to make some changes to the terms with which the following discussion is framed, but I am confident that the changes would not significantly effect the conclusions of those discussions.

2.3 The Challenge to the Unity Thesis

In the introduction, I suggested that the verdict delivered by introspection and reflection on one's own experience is that consciousness is phenomenally unified. No doubt only a relatively small number of people engage in the kind of reflection on their

own experiences that would lead them to entertain the question regarding the extent to which consciousness is unified, but, to the few who do, the verdict they reach is likely to be consistent with the Unity Thesis. The Unity Thesis carries a high degree of *prima facie* plausibility. Perhaps the modal operator in the thesis lowers its initial plausibility somewhat, but those who deem their own experience to be phenomenally unified are likely to believe that this aspect of their own experiential lives generalizes at least to some extent. They may not be confident that it generalizes as broadly as the modal operator implies it does, but they are likely to believe that their own experience is not unique in this regard. One reason for this is that it is very difficult to imagine what it would be like to have an experience that was not phenomenally unified. In this section, I will introduce the biggest challenge to the Unity Thesis: the empirical results that have been gathered from studying split-brain subjects.

2.3.1 Split Brains

In the middle of the previous century, a last resort treatment for patients suffering from severe epileptic seizures was introduced. It consisted of surgically severing the corpus callosum, the thick strand of nerves that connects the two hemispheres of the upper brain. The surgery was first attempted in 1940 by van Wagenen and Herren after they noticed that the rate of seizures experienced by epilepsy subjects who had developed tumors in their corpus callosum decreased as the tumors developed. They performed the surgery on a number of subjects with significant success. Most of them experienced fewer and less severe seizures after the surgery. In the 1950s, Sperry and his colleagues began to study the effects of the surgery in animals and then, in the 60s, they adapted their techniques for human subjects. Sperry was eventually awarded the Nobel Prize for his work, and much of what we know today about the relative

specializations of the two hemispheres of the brain we know as a direct or indirect result of the pioneering research conducted by him and his colleagues.¹¹

Given the seemingly invasive and severe nature of the surgery, one would expect patients undergoing this treatment to be significantly compromised as a result. Surprisingly, this is not the case. After recovering from the surgery, a process that often takes a year or more, patients are able to carry on a remarkably normal existence. Sperry reported that a commissurotomy patient that survived the surgery without complications “could easily complete a routine medical examination without revealing to an uninformed practitioner that anything [was] abnormal. Nor [was] there any marked change in verbal scores on the standard IQ test” (1984, 662). This surprising result is what led him and his colleagues to study these subjects more thoroughly. What they found, when they placed split-brain subjects in certain controlled circumstances, is that there are some significant ways in which they differ from normal subjects.

In a typical scenario, split-brain subjects were seated at a table with their hands and the surface of the table shielded from view by a screen. If an object was placed in one hand, split-brain subjects were unable to pick the same object out of a pile of objects with the other hand. When the subjects were given an object to hold in each hand, they could not say whether the objects were the same or different. Sometimes subjects were asked to focus their gaze on a fixation point on the centre of a screen. Words or pictures of ordinary household objects were then flashed with a tachistoscope on to either the left or the right half of the screen. When presented with a word or picture in one half of their visual field in this way, split-brain subjects could not recognize the object when it was presented on the opposite side. If an object was

¹¹For some helpful and brief summaries of split-brain research, see Mathews et al. (2008) and Gazzaniga (2005). For a more comprehensive survey of the data gathered from split-brain research and an account of the specializations of the two hemispheres, see Gazzaniga (2000).

shown in the left visual field, the subjects could identify the object manually with their left hand but could not identify it verbally. If an object was shown in the right visual field, the subjects could identify it verbally but not manually with their left hand. Split-brain subjects were also tested in other ways. In other trials, the subjects could not tell whether a line was continuous or broken if the break in the line coincided with the divide between the two halves of their visual field.

For the most part, the right hemisphere receives sensory information from the left side of the body, whereas the left hemisphere receives sensory information from the right side of the body. The right hemisphere receives visual information from the left half of the visual field, and the left hemisphere receives information from the right half of the visual field. The same holds with respect to auditory information. There are a few sensory systems that feed information to both halves of the upper brain; both hemispheres, for instance, are attuned to both sides of the face and to those systems that detect pressure on the skin and position of the limbs. The movements of the limbs are, for the most part, managed contralaterally, and the two hemispheres have different specializations.¹² In most people, the left hemisphere is dominant for language, whereas the right hemisphere is dominant for visuo-spatial processing. One of the ways in which the right hemisphere's area of expertise is manifest is in the dominant role it plays in recognizing faces. And although the left hemisphere is usually dominant for language, the right hemisphere is often responsible for the ability to sing (Bogen and Gordon, 1971).

The experiments run on split-brain subjects are designed to exploit these neuro-anatomical facts, and what the results of these studies seem to show is that sensory

¹²This is relatively common knowledge now and was, to some extent, common knowledge in the early days of Sperry's research. However, as has been suggested above, some of what we know about the relative specializations of the two hemispheres is knowledge that was acquired through split-brain research. See Levy (1990) and Gazzaniga (2005) for a discussions of the contributions of split-brain research in this area.

information that is fed to one hemisphere is not available to the other. This has led researchers and theorists to make some rather startling claims. Sperry, for instance, is on record as saying that “[e]ach hemisphere can be shown to experience its own private sensations, percepts, thoughts and memories that are inaccessible to awareness in the other hemisphere. . . . [E]ach surgically disconnected hemisphere appears to have a mind of its own, each capable of controlling the behaviour of the body but each cut off from, and oblivious of, conscious events in the partner hemisphere” (1984, 663) and that split-brain subjects “have, in effect, not one inner visual world...like the rest of us, but rather two separate and independent inner visual worlds—one for the right and one for the left half field of vision, each in their separate hemispheres” (1968b, 306). The split-brain research was introduced into the philosophical literature by Nagel (1971). As Nagel rightly realized, and as the above comments by Sperry make clear, the observations gained from the study of split-brain subjects pose a sort of puzzle for theorists interested in understanding the nature and structure of consciousness, and, depending on how the the data are ultimately interpreted, have significant implications for our understanding of personhood.

Nagel asked the following question: how many minds do split-brain subjects have? The idea that an individual could have more than one mind seems absurd. The challenge for a view that is more consistent with our naïve folk psychology, however, seems a little strained when we try to reconcile it with the empirical data. How could a single mind be both aware of an object, as demonstrated by its ability to identify the object with, say, the left hand, and unaware of the same object, as demonstrated by its inability to identify it verbally? We could defend our naïve views by maintaining that activity in the right hemisphere, the hemisphere that was historically known as the silent or minor hemisphere, is unconscious and that, when it identifies an object shown in the left visual field, it does so unconsciously. But this seems somewhat *ad hoc*. The right hemisphere is, at least on a macro-scale, relatively anatomically

isomorphic to the left hemisphere and is, on a micro-scale, just as complex; it houses as much neural activity as is housed by the left hemisphere. Why should the left hemisphere be capable of supporting consciousness and the right hemisphere not?

Sperry's speculation, according to which these subjects have two minds, one for each hemisphere, seems more consistent with the data. The mind associated with the right hemisphere consciously represents and responds to visual information from the left visual field, whereas the mind associated with the left hemisphere consciously represents and responds to visual information from the right visual field. The two minds are distinct, and there is no direct exchange of information between them. That is why split-brain subjects are able to manually identify objects presented in their left visual field but unable to verbally identify them, and that is why there is a similar disparity between how the two hemispheres respond to objects presented in the right visual field.

This view also seems to be in a particularly good position to account for some observed cases in which the two hemispheres behave in overtly competing ways. Mark (1996, 192) reports the following experience with a split-brain patient he was treating.

I asked the patient whether she were still having seizures. Again she gave contradictory responses and appeared frustrated. She then tried to enumerate her seizures. The right hand held up two fingers, but then the left hand reached over and tried to force it down. After trying several times to tally her seizures—and failing each time—she paused, softly counted to ten, and then simultaneously displayed three fingers with her right hand and one with her left. She appeared astonished and confused and asked why she behaved this way.

Wilkes (1978, 187) gives the following account of a scene recorded by Sperry.

The [split-brain subject] was asked to form a structure from a number of building blocks, using the right hand only; this was done in full view.

The right hand was clumsy and slow, and the left hand moved in to help, brushing aside the right hand. The experimenter pushed away the left hand. A few moments later the left hand again attempted to take over; this time the patient, with his right hand, grasped the left wrist and pushed it away. But you cannot keep a good hand down; in the end the patient had to sit on the intrusive left hand, to prevent it from interfering further.

In another oft-cited case, a split-brain subject was given a pipe to hold in his left hand out of sight. After the pipe was removed, he was asked to write with his left hand what he had just held. He slowly wrote “P” and then “I.” Then, in an abrupt change of pace, the “I” was made into an “E” and the word “PENCIL” quickly completed. Shortly after the word “PENCIL” was completed, the last five letters were crossed out and the hand drew a crude drawing of a pipe.¹³ One explanation of this case is that the right hemisphere knew what the left hand held out of sight, but, because it is not as adept at language as the left hemisphere, struggled to identify in writing what the left hand held. The left hemisphere observed what was being written on the page and, when it decided what the hand was attempting to write took over control of the hand and wrote “PENCIL.” The right hemisphere, seeing the incorrect response on the page took over the hand again and made the correction. If the cases in which subjects are merely unable to identify objects that have been shown to the opposite hemisphere do not adequately motivate the view that split-brain subjects have two minds, these cases in which subjects behave in explicitly conflicting ways surely go much further. In these cases there seem to be two distinct seats of action-generating behaviour that are struggling with each other. One hemisphere responds to a question in a way that is consistent with what it knows only to have the other

¹³Wilkes (1978) and Nagel (1971) attribute this report to Levy (1969).

hemisphere interfere and assert itself in a conflicting manner. How else, other than positing two independent and distinct centres of consciousness, are these data to be explained?

Although positing two minds in split-brain subjects seems to make it easier to account for some of these observations, there are other considerations that make it much less attractive. If these subjects have two minds, then how is it that they can behave like normal subjects outside of experimental conditions? Should they not exhibit conflicting behaviour outside of experimental conditions just like they do when they are in them? And then there is the way in which maintaining that split-brain subjects have two minds pushes squarely and strongly against our naïve folk psychology. Could it really be that these subjects have more than one mind? If they have more than one mind, do they always have two minds or does a second mind pop into existence at the onset of the experimental conditions and go out of existence once the experiments have been completed? The view that a second mind pops into existence at the commencement of the experimental conditions strains credulity even more than the view that these subjects have two minds running in tandem constantly. Although there is a significant change to brain structure that occurs for these patients at the time of the commissurotomy, the experimental conditions do not tamper with the brain's physical structure at all. All that happens in these experimental conditions is that the subject is shown various stimuli in various controlled ways. How could a conscious mind pop into existence merely in response to various external stimuli? It seems that if we go the two-minds route, we need to say that these subjects have two minds running in tandem all the time, not just in the experimental conditions. But if we say this, are we then committed to saying that normal subjects, subjects that have intact corpus callosums such as ourselves, also have two minds running in tandem constantly? If the two hemispheres are each capable of supporting a conscious mind after a commissurotomy, why not before?

2.3.2 A Range of Views

Nagel framed the puzzle generated by the split-brain data as a question about how many minds split-brain subjects have. His answer was that it is indeterminate how many minds split-brain subjects have. Although Nagel can be credited with initiating philosophical discussion about split-brain subjects and the structure of their inner experiential lives, few have defended his view. Most have found the idea that it is *indeterminate* how many minds split-brain subjects have to be implausible, and, instead of framing the discussion as a discussion about how many minds split-brain subjects have, most have framed the debate as a question about whether or not, or to what extent, the consciousness of these individuals is (phenomenally) unified.

One view that has received a significant amount attention but few adherents is Lockwood's (1989; 1994) view that the consciousness of split-brain subjects is partially unified. Roughly, the view is that split-brain subjects have two streams of consciousness that overlap. In Lockwood's more technical vocabulary, the central tenet of the view is that the relation of *co-consciousness* is not transitive, making possible a state of affairs, instantiated by split-brain subjects, in which *phenomenal perspectives* partially overlap.¹⁴

Other theorists maintain views that are closer to the view suggested by Sperry's comments quoted above. These theorists agree with Lockwood in maintaining that split-brain subjects have two streams of consciousness, but they reject Lockwood's claim that the streams partially overlap, insisting, instead, that the two streams of consciousness are independent from each other and (completely) disjoint. Amongst these theorists, the default view is that the consciousness of split-brain subjects becomes (and remains) split at the time of the commissurotomy (see Koch, 2004), but there are some views that diverge from this default. For instance, Puccetti (1981)

¹⁴Baumann (2007) defends a view that may be similar to Lockwood's in that he argues for the claim that phenomenal unity can be a matter of degree.

argues that the surgery does not introduce a split in consciousness; even normal subjects with an intact corpus callosum have a split consciousness. What happens as a result of the commissurotomy is that the two streams of consciousness become limited with respect to the sensory information they each receive. Others, like Marks (1981) and Tye (2003), maintain that the consciousness of split-brain subjects only splits into two separate streams during exceptional circumstances like those that obtain in the experimental conditions.

Finally, there are some who reject the claim that split-brain subjects have two (overlapping or non-overlapping) streams of consciousness. They maintain that split-brain subjects, like normal subjects, have but one stream of consciousness. Bayne (2008, 2010) has offered, perhaps, the most significant defence of this view. Although it may be a divergence from his earlier position (see Gazzaniga and LeDoux, 1978), Gazzaniga (2000) also seems to accept a one-stream view.

Of these different interpretations of the split-brain data, one-stream views are most amenable to the Unity Thesis and to our pre-theoretic and naïve ideas about the nature and structure of conscious experience. But being amenable to our naïve judgements about the nature and structure of conscious experience is only one consideration that is in play in the shaping of our all-things-considered view on the matter. We must also take the findings of empirical research into account. This is where the waters become muddied. In many ways (though not all), two-stream accounts are better positioned to accommodate the split-brain data. On the other hand, they challenge our preconceived ideas about conscious experience. In the next two chapters, I will attempt to navigate these muddy waters. I believe that our pre-theoretic and naïve view about the nature and structure of conscious experience (at least with respect to the (phenomenal) *unity* of consciousness) happens to be correct and I believe that it can be defended. In chapter 3, I will consider and argue against two-stream interpretations of the split-brain data. One-stream views will be the topic of chapter

4. The main challenge for one-stream views is reconciling the view with the split-brain data. Bayne attempts this reconciliation by defending an interpretation of the split-brain data according to which the two hemispheres of the upper-brain take turns supporting a single stream of consciousness. Like Bayne, I am inclined to accept the Unity Thesis and a one-stream interpretation of the split-brain data, but I believe Bayne's "switch model" is problematic. I will argue against Bayne's switch model and present my own interpretation of the split-brain data.

CHAPTER 3

TWO STREAMS OF CONSCIOUSNESS

What I want to do in this chapter is argue against the different two-stream views presented at the end of the previous chapter. In order to do so, however, it will be necessary to have a more precise understanding of the views themselves. This will require, in turn, a more precise understanding of some of the key notions involved in the formulations of these views. In the first section of this chapter, I will unpack these notions. In the rest of the chapter, I will argue that the different two-stream views should not be accepted.

3.1 Maximal States of Consciousness

The notion of a stream of consciousness is part of our naïve folk psychology. Our conscious inner lives unfold over time. Although I am having a certain set of experiences at the moment, the experiences I have change as I move about and interact with the external world. Because our experiential lives are always changing, always in flux, it is natural to think of them as being like rivers or streams. One cannot, as they say, step into the same river twice. And just as a stream extends both upstream and downstream from one's particular vantage point on its bank, our experiential lives have temporal extension. Although the expression "stream of consciousness" is, perhaps, used primarily to refer to the ever-changing flux of experience in the ever-changing now, it can just as well be used to refer to one's experiential life in its entirety, i.e., as the collection of experiences of which one's past, current, and future experiences are all parts.

When I introduced the notion of phenomenal unity in section 2.1, I said that experiences are taken to be phenomenally unified when they are part of the same global or encompassing experience. Intuitively, a maximal or global state of consciousness is just the entirety of one's experience at a time. It is a temporal slice of one's stream of consciousness and is a state of consciousness that encompasses, or has as parts, all of one's conscious states at that time. Although the notions of a stream of consciousness and of a maximal or global state of consciousness are intuitively clear and are, for the most part, thought of in these ways, these characterizations are not quite adequate in the present context. The reason for this is that some of the theorists in the debate over how to understand the inner lives of split-brain patients take issue with this way of characterizing these notions. One way to characterize the debate is to say that it concerns the answer to the following question: does a maximal state of consciousness always include all of the experiences its subject has at that time? If we take the notion of a maximal state of consciousness to be the notion of a state that consists of all of a subject's conscious states at a time, then the answer to this question is affirmative and a straight-forward entailment from the notion itself. But not all of the theorists in the debate accept an affirmative answer to this question. Since there is disagreement about this, and since the different views presented at the end of the previous chapter employ the notion of a maximal or global state of consciousness, we will need to characterize this notion in such a way that the different views can be clearly delineated without begging any questions. This will be the project in the remainder of this section. I will arrive at a consensus characterization somewhat indirectly. Instead of stating what I will call the "Consensus View" at the outset, I will spend some time motivating it independently. This will be useful in that some of my objections to the different two-stream views that I will raise in section 3.2 will rely on some of the considerations I will use to motivate the Consensus View.

3.1.1 Subjects, Conjoint Phenomenology, and Maximal States

I begin by making an assumption. I assume that the arguments in section 2.2.3 are sound and that it is both meaningful and coherent to talk about a conscious state being *part* of another conscious state. As I indicated there, the notion of parthood involved is nothing more than the formal notion of parthood in classical mereology. If it is meaningful to talk about a conscious state being part of another, then it is also meaningful to talk about (mereological) *sums* of conscious states. And if it is meaningful to talk about sums of conscious states, then saying what it is to be a maximal state of consciousness will involve answering the following question: which sums of conscious states are maximal states of consciousness?

Given the way in which the notion of a maximal state of consciousness was intuitively presented, i.e., as a temporal slice of a stream of consciousness, the first thing that one might say about maximal states of consciousness is the following.

A sum of conscious states, u , is a maximal state of consciousness only if
there is a time, t , such that every part of u occurs at t . (3.1)

Less formally, this thesis says nothing more than that all the parts of a maximal state of consciousness must occur simultaneously. Although this thesis may seem to flirt with triviality, the issue is complicated somewhat by the possibility of temporally extended experiences. There is, for instance, a live debate about what one should say about our consciousness of certain temporally extended events like melodies in musical works. A melody, perhaps by definition, has temporal duration. How else could a melody be distinguished from the sounding of a single note? So it seems there is a legitimate sense in which our experiences of musical works are temporally extended. One might argue that a maximal state of consciousness ought to be the kind of thing that could include within itself a conscious experience of a temporally extended event

like a melody. On this kind of view, it would be problematic to understand maximal states of consciousness as states that do not have temporal extension.

Although the correct way to understand temporally extended experiences is a legitimate issue, I believe it is an issue that can be passed over in this context. Temporally extended experiences often temporally overlap without being contemporaneous throughout their respective durations. For example, when listening to an organ work by J. S. Bach, one statement of the fugue subject often begins before another has ended completely. If my (maximal) experience while both statements of the fugue are sounding consists of both statements of the fugue subject and my experiences of them are both temporally extended, then not every part of these two experiences can be simultaneous with each other. Otherwise my experience would be of the two fugue statements sounding in unison rather than in canon. Even if it should turn out to be problematic to regard temporally extended experiences as having temporal parts, our final analysis of experience, even, and perhaps especially, temporally extended experience, must have some way to refer to the temporal *stages* of such experiences. Whatever notion is used in such a final analysis to refer to the temporal stages of experience could be used in a reformulation of thesis (3.1). Thus, I will take thesis (3.1) to express a truth about maximal states of consciousness.

Another thesis about maximal states of consciousness that may seem to flirt with triviality is the following thesis about the subjects of maximal states of consciousness and their parts.

A sum of conscious states, u , is a maximal state of consciousness only if there is a subject of conscious states, s , such that s is the subject of every part of u . (3.2)

Less formally, this thesis says that all of the parts of a maximal state of consciousness must have a common subject. Thesis (3.2) makes having a common subject

a necessary condition on conscious states that compose a maximal conscious state. There are several things to say about this thesis. First, the notion of a subject that it invokes is that of a mere subject of consciousness. Second, one might wonder whether the condition is not also sufficient, i.e., whether the conditional in (3.2) should be a bi-conditional. The problem with making the thesis a bi-conditional is that it would then be inconsistent with (3.1). Subjects of experience have experiences at different times. Thus, there are sums of experiences whose parts all have a common subject but do not occur simultaneously.

Third, although (3.2) may seem to be an obvious truth about maximal states of consciousness, it makes an assumption that not all theorists accept, namely, that conscious mental states (or experiences) have subjects. In a famous passage, David Hume noted that all that introspection ever provides access to is experience; one never stumbles across oneself at a moment when one is not having an experience.¹ One never encounters a self or subject that is the owner of experience. Some, perhaps Hume himself, have taken this observation as a reason to believe that there are bundles of experiences or conscious states, but that, properly speaking, there are no such things as subjects of those experiences.² Proponents of this view, even if they are inclined to countenance the phenomenon of phenomenal unity and provide an explanation of it, will object to (3.2). I will not take up the bundle theory here. Although the discussion here and in the following pages will proceed on the assumption that experiences have subjects, and considerations about subjects of experience will often influence the direction the arguments will take, the ultimate question is about how many streams of consciousness split-brain subjects have. This is a question that must be addressed by bundle theorists as well as by those who countenance subjects of experience. Of

¹See *A Treatise of Human Nature* (section 1.4.5).

²The qualification “properly speaking” is important for it is open to a proponent of such a view to define subjects *in terms of* such bundles.

course, bundle theorists will not employ the notion of a subject of experience in their account of what a maximal state is, but they will most likely employ the notion of conjoint phenomenology, which will be introduced below. Indeed, it is far from clear that they there is another notion at their disposal that could be used to define maximal states. Once bundle theorists have settled on an account of a maximal state of consciousness, some of the arguments I will raise below will proceed just as well on their account of a maximal state.

Not everyone will accept (3.1) and (3.2), but I take (3.1) to be a relatively innocuous claim about conscious states and (3.2) enjoys the status, at least with respect to its assumption that conscious states having subjects, of being orthodox. In my view, these two theses give us a pair of truths about maximal states of consciousness. Moreover, it seems to me that, aside from Tye, whose view that experiences do not have parts may preclude him from accepting (3.1), the theorists I will be engaging in the following pages would also be inclined to accept them. Given their plausibility and their relatively widespread acceptance, one might think that these two theses, taken together, are enough to characterize maximal states of consciousness. Thus, one might be tempted to accept the following.

A sum of conscious states, u , is a maximal state of consciousness if and only if (i) there is a time, t , such that every part of u occurs at t and (ii) there is a subject of conscious states, s , such that s is the subject of every part of u . (3.3)

Thesis (3.3) goes a long way toward capturing our intuitive notion of what a maximal state of consciousness is, but it does not go quite far enough. The reason is that it is much too permissive; states of consciousness that are too “small” to be maximal states of consciousness turn out to be maximal according to (3.3). Take any proper subset of your current states of consciousness. All the states in that subset will be

simultaneous and had by the same subject, namely, you. Thus, any proper subset of one's current state of consciousness will satisfy the necessary and sufficient conditions laid out by (3.3), and so, if it is accepted, then one is committed to saying that we have many more maximal states of consciousness than we believe ourselves to have.

The most obvious way to deal with this problem would be to add a clause to (3.3) to the effect that a sum of conscious states is a maximal state only if it includes all of a subject's conscious states at a time.

The Naïve View A sum of conscious states, u , is a maximal state of consciousness if and only if (i) there is a time, t , such that every part of u occurs at t , (ii) there is a subject of conscious states, s , such that s is the subject of every part of u , and (iii) all of s 's conscious states at t are in u .

As I mentioned at the outset of this section, the intuitive notion of a maximal state is of a state that consists of all of a subject's conscious states at a time. The Naïve View captures this intuitive sense well. It says that a sum consisting of all of one's conscious states at a time is a maximal state and it entails that sums consisting of only a proper subset of those states are not maximal states. I also pointed out at the beginning of this section that the intuitive notion of a maximal state is of a temporal slice of a stream of consciousness. The Naïve View would seem to have the consequence that subjects of consciousness can have only one stream of consciousness at a time. If the Naïve View were to be accepted as our working characterization of maximal states of consciousness, several interpretations of the split-brain data would be ruled out prematurely. Thus, an alternative characterization of maximal states of consciousness must be formulated.

Fortunately, we do not have to look too far to find one. The states that compose a maximal state have another property that the Naïve View makes no mention of: they are all experienced together in the way that is indicative of phenomenal unity; they have a conjoint phenomenology. If my visual experience of my mug and my

auditory experience of music are both part of a maximal state of consciousness, then the phenomenal character of my maximal state includes both the visual experience of the mug and the auditory experience of the music; the visual experience of the mug and the auditory experience of the music are experienced together. This can be formalized as follows.

A sum of conscious mental states, u , is a maximal state of consciousness only if, for any pair of conscious states, c_1 and c_2 , that are (proper or improper) parts of u , c_1 and c_2 have a conjoint phenomenology. (3.4)

There are a couple of things that should be noted about thesis (3.4). First, it is assumed that, if c_1 happens to be a proper part of c_2 , there is nothing more to the what-it's-like-ness of having c_1 and c_2 together than there is to the having of c_2 itself. Of course, there is more to the having of c_2 than there is to the having of c_1 — c_1 is a “smaller” state than c_2 —but the what-it's-like-ness of c_2 includes the what-it's-like-ness of c_1 , so the having of c_1 in “addition” to the having of c_2 adds nothing to what it is like for a subject to have them both. Thus, there is nothing particularly mysterious about the conjoint phenomenology of a state of consciousness that happens to be a maximal state and a state that is one of its proper parts.

Second, although thesis (3.4) has been presented as a thesis that is independent from (3.1) and (3.2), given what it means for a pair of conscious states to have a conjoint phenomenology, (3.4) may well *entail* (3.1) and (3.2). For instance, it is unclear how a pair of conscious states could have a conjoint phenomenology without being simultaneous and had by the same subject. Few theorists draw attention to this entailment explicitly, but it seems to be implicitly recognized by the conceptual priority that they give to (their respective analogues of) conjoint phenomenology within their conceptual frameworks.

All that is left to do is make the appropriate adjustments to the Naïve View. We need to remove clause (iii) and insert the condition outlined in (3.4). But we also need to add an additional clause. If all we did was substitute the condition regarding conjoint phenomenology for clause (iii) we would be left with the same problem that plagued (3.3): the resulting view would have the result that states that are too “small” to be maximal states of conscious would be maximal states. The reason for this is that every proper subset of a sum of states that satisfies the necessary condition in (3.4) will also satisfy the necessary condition. Thus, we also need a clause that will ensure that the only sums of conscious states that turn out to be maximal states are the biggest states that satisfy the necessary condition in (3.4). The resulting thesis is somewhat cumbersome, but I suggest the following.

The Consensus View A sum of conscious states, u_1 , is a maximal state of consciousness if and only if (i) there is a time, t , such that every part of u_1 occurs at t , (ii) there is a subject of conscious states, s , such that s is the subject of every part of u_1 , (iii) for any pair of conscious states, c_1 and c_2 , that are (proper or improper) parts of u_1 , c_1 and c_2 have a conjoint phenomenology, and (iv) there is no sum of conscious states, u_2 , such that u_1 is a proper part of u_2 and u_2 satisfies all of the conditions stated in (i)–(iii).³

The main difference between the Naïve View and the Consensus View, the difference that will be important in the following pages, is that the Consensus View allows for the possibility that the subject of a maximal state of consciousness has conscious states that are contemporaneous with the maximal state but not part of it. This is a possibility that runs counter to our naïve and intuitive view about what maximal states of consciousness are, and the reasons why theorists have required this possibility to be left open will be discussed in more detail shortly, but because it is left open,

³Of course, in order to satisfy the conditions stated in (i)–(iii), “ u_1 ” would have to be substituted with “ u_2 .”

the Consensus View is one that all theorists who weigh in on phenomenal unity and the split-brain data can accept. Thus, it will be our working conception of a maximal state of consciousness. Before moving on to consider the views of these theorists, however, it will be worth pointing out that there are some additional unorthodox possibilities that the Consensus View leaves open. These possibilities will be the topic of the next two sections, and will refer to these possibilities as the possibility of a “collective consciousness” and “spread consciousness” respectively.

3.1.2 Collective Consciousness

Consider the sum of conscious states that consists of all of my current conscious states and all of your current conscious states. Call it “JOINT.” Now consider the following question: is JOINT a maximal state of consciousness? The immediate response is to say that it is not. I believe this is the correct answer to the question, but the Consensus View does not, by itself, deliver this verdict, at least not in any straightforward or obvious way. The immediate response is rooted in the following line of thought: I know that I am the subject of half of the conscious states that compose JOINT and I know that I am not the subject of the other half of the conscious states that compose it; thus, there is no (single) subject that has all of the conscious states that compose JOINT.

As natural as this line of thought is, there are two ways that it might be called into question. The first involves questioning its validity. The mere fact that *I* am not the subject of half of the conscious states that compose JOINT does not straightforwardly entail that there is *no* subject that has all of the conscious states in it. People sometimes talk about a collective consciousness. It is seldom clear what people have in mind when they talk about it, but, when the topic of a collective consciousness comes up, people seem to use the expression to refer to a kind of conscious subject that somehow transcends that of the individual. This is still rather vague, but suppose

that a collective subject of consciousness is one that can enjoy the conscious states of many different individuals all at once. If there are such subjects of consciousness, then there might be one that enjoys all of the states that compose JOINT, in which case JOINT would satisfy the conditions laid out in the Consensus View and be a maximal state.

It is hard to take the notion of a collective subject seriously. In order to see why, it must be made clear exactly what is being suggested when such a subject of consciousness is posited. A collective subject does not merely have some conscious states that are *qualitatively* identical to those of the lesser subjects it shares those states with. Rather, a collective subject is the subject of the very same conscious particulars that the lesser subjects are subjects of; a collective subject's conscious states are not merely type-identical to the conscious states of the lesser subjects but token-identical.

This distinction between type- and token-identical states is important. Although the likelihood of there actually being a pair of subjects with (exactly) type-identical conscious states at a time is extremely low, it is not too difficult to describe a set of circumstances in which there might be such a type-identity. White (1987), for instance, describes a “group fusion” scenario in which the sensory systems of a number of human subjects are wired so that they all receive a single set of sensory inputs from a single body. The subjects are also wired so that each of their respective pre-conscious intentions are fed into a mediation device. This device selects one set of pre-conscious intentions and sends it back to the subjects, ensuring that, when the pre-conscious intentions become conscious in the different subjects, each of them will have the very same conscious intentions.⁴ Their respective, functionally indistinguishable conscious intentions are then sent to the motor systems of the body they are all

⁴There is a real question about whether the respective *contents* of these different conscious intentions are identical, but the conscious intentions will have a qualitatively identical feel subjectively.

receiving sensory information from. There are many other analogous examples in the literature. Although these cases are often somewhat of a science-fiction variety, they seem to be both metaphysically and nomologically possible. What is important for this discussion here is that it is not too much of a stretch to believe that the different human subjects in White's group-fusion scenario have type-identical conscious states.

By contrast, it is very difficult to wrap one's head around a situation in which two subjects have *token*-identical experiences. One way to conceive of a situation like this is to think of an experience as being like a movie shown in a theatre. In a theatre, there is one moving image on a screen that can be viewed by many people at once. There is a danger, however, with thinking of experiences in this way: it lends itself to thinking of the relation between the subject and experience as being analogous to the relation between a perceiver and a perceived object. Once experience is thought of as an object of perception in this way, questions can be asked about how the subject of the experience perceives experience. Does the perception of the experience itself involve the having of an experience? In order to be consistent, the defender of this model seems compelled to answer in the affirmative. And once this happens, it seems the defender of the theatre model of experience is committed to an infinite regress of experiences.⁵

One way for the defender of the theatre model to block this consequence is to argue that, although experiences can be enjoyed by multiple subjects in the same way that external objects can be perceived by many perceivers at once, the relation between subject and experience is not the relation of perception. Rather, the relation is a *sui generis* appearance relation that does not require the having of a second experience in order to have the first, so no infinite regress of experiences is entailed by the existence of one.

⁵See Dennett (1991) for some more criticisms of the theatre model.

Sui generis relations are often met with suspicion. The reason for this is that they are often posited at the end of analysis to avoid some problematic consequence. This can make them seem *ad hoc*. Also, given their *sui generis* nature, they resist being understood in terms of more familiar and less mysterious relations. Often this contributes to the sense that the posited *sui generis* relation comes with a theoretical cost. It is questionable, however, whether the defender of the theatre model does anything untoward by positing such a relation. The relation between a subject and experience is a mysterious relation on any view. Thus, if positing a *sui generis* relation is a problem for the defender of the theatre model, it is a problem shared with many others.⁶

This is not to say that the theatre model, or any view that posits the possibility of multiple subjects of experience, is without other problems. For one, such views lead to some awkward questions. Some of these awkward questions have already been encountered but there are others. For instance, once it is allowed that conscious states can literally be shared by more than one subject of consciousness, then one can ask what determines how many subjects a given conscious state has. This question requires an answer. If it is not answered, then there is no guard against a rather unpalatable ontological inflation. For example, if there is no answer to this question—and here I mean not just that we do not *know* what the answer is but that there is no answer to be found—then there would be no reason to rule out the possibility that there are an indefinite number of subjects of consciousness that share my conscious states with me.

One might attempt to curb the ontological inflation by saying that subjects of experience are to be individuated by the conscious states they have. This would rule out there being multiple subjects with the very same set of conscious states, but it

⁶Most others, however, would argue, contrary to the defender of the theatre model, that this *sui generis* relation does not hold between a single experience and multiple subjects.

would not rule out there being distinct subjects of consciousness whose conscious states overlap to a large degree. Subjects of experience would be enumerated by enumerating the number of distinct sums of experiences there are at a time. This would curb the ontological bloat somewhat, for it would rule out the possibility of there being more than one subject for the same set of conscious states, but it is not clear that the result is any more palatable than the previous one. At the end of the day, the notion of a subject of consciousness we would be left with would make subjects of consciousness about as insignificant and uninteresting as arbitrary sums of experiences are. What we are interested in, when it comes to subjects of experience, is how they are related to agency and intentional action. If there is a subject of experience for every arbitrary sum of experiences, then the notion of a subject of experience would become irrelevant to such topics of enquiry.

The fact that a thesis has implausible or unorthodox implications is often not decisive against it. Nevertheless, I believe that considerations like these are reason enough to believe that conscious states cannot have more than one subject and that the following is a necessary truth.

For any pair of subjects (of conscious states), s_1 and s_2 , and any conscious state, c , if s_1 is a subject of c and s_2 is a subject of c , then s_1 and s_2 are identical. (3.5)

This is hardly a minority view, so it is certainly in play and available for use in deliberations about whether JOINT is a maximal state of consciousness. It is also worth noting that none of the theorists in the debate over split-brains and phenomenal unity explicitly or implicitly reject it. But conjoining (3.5) with the Consensus View doesn't yet rule out JOINT as a maximal state of consciousness.

3.1.3 Spread Consciousness

There is a second way the line of reasoning behind the immediate response can be challenged. Above we examined a kind of conscious subject that could literally share the conscious states of any number of distinct conscious subjects. If there are such subjects of consciousness then a state of consciousness like *JOINT* could involve three distinct subjects; I would have half of the states in *JOINT*, you would have half of them, and the collective subject would have them all. Another way to challenge the line of reasoning behind the immediate response is to suggest that *JOINT* has only one subject of consciousness, a subject of consciousness that has all of the conscious states that compose *JOINT* and does not share them with any other subject. Positing such a subject of consciousness would challenge the line of reasoning behind the immediate response by questioning its premise; when I think to myself that I am the subject of half of the states in *JOINT*, I am correct, but I am mistaken when I think to myself that I am *not* the subject of the other half of the states in *JOINT*. If, in thinking to myself, “I am the subject of half of the states in *JOINT*,” I am successful in referring to a subject of consciousness with the first-person reflexive pronoun “I,” then I refer to a subject of consciousness that is “bigger” than I think it is; I refer to a “spread” subject of consciousness that has both the states I believe myself to have and the states you believe yourself to have. This view, in essence, erases the boundaries between subjects of experience in the traditional sense. You and I are not distinct subjects of experience. Rather, we together make single subject of experience that has all of the states in *JOINT*.

Positing this kind of collective or spread subject of consciousness does not commit one to a denial of thesis (3.5), but it flies in the face of our subjective sense that we are distinct subjects of consciousness from each other. Should we take the notion of a spread subject of consciousness seriously? I believe we should not.

The fact that the notion of a spread subject of consciousness flies in the face of our ordinary and intuitive sense of what a subject of consciousness is is significant, but, again, more argument is needed. One could try to argue against the notion of a spread subject in the following way. If a mental state is a conscious one, it must be conscious *for* a subject. A conscious state, perhaps by definition, is one that has a phenomenal character, a what-it's-like-ness. If a state has a what-it's-like-ness, then there must be some subject of consciousness for whom there is something it is like to be in it. But for half of the states that compose JOINT, there is nothing it is like for me to be in them; if they have a phenomenal character, I am not privy to it. This would seem to imply that I am not the subject of those states. If they have a subject, it must be distinct from me. Thus, there is no subject that has all of the states in JOINT together.

I believe this to be a sound argument, but it is worth considering how a defender of the notion of a spread subject might respond. Its defender is likely to maintain that the subjective facts invoked in the argument are unsurprising and perhaps to be expected when there are spread subjects of consciousness. A spread subject, its defender might argue, can have parts just like an experience can have parts. The mere fact that there is something it is like for a spread subject to be in a certain conscious state does not entail that every part or aspect of that subject enjoys that state's what-it's-like-ness. One part of a spread subject may well be oblivious to some of the conscious states that the spread subject as a whole has. One may well wonder what this talk of parts of a subject amounts to. One might also wonder whether the notion of a part of a subject is just a disguised way of maintaining that experiences can have more than one subject. However, if the notion of a spread subject is coherent—I do not take myself to have demonstrated that this is the case—and if it can be made plausible, then perhaps it is possible for JOINT to satisfy condition (ii) of the Consensus View.

Another question remains: would a situation in which JOINT was had by a spread subject be a situation in which JOINT satisfies condition (iii) of the Consensus View, the condition involving conjoint phenomenology? The *prima facie* answer would again seem to be negative. The backing for this answer would come from the same kind of considerations that were just raised against the notion of a spread subject. For half of the states in JOINT, there is not anything it is like for me to be in them. If I am not privy to a state's what-it's-like-ness, if a state does not have an occurrent phenomenal character for me, then it would seem impossible for its phenomenal character, if it has one, to be experienced together with that of another state. There being something it is like for me to be in a state would seem to be a prerequisite for that state to have a conjoint phenomenology with any other state. Since I am not privy to the phenomenal character of half of the states in JOINT, it would seem that JOINT would fail to satisfy condition (iii) of the Consensus View.

Here again a defender of the notion of a spread subject is likely to reply by pointing out that this argument is offered from the perspective of one of the parts or aspects of the spread subject. The mere fact that one part or aspect of the spread subject does not enjoy the phenomenal character of half of the states in JOINT does not entail that the subject as a whole does not enjoy all of the states in JOINT. And if it is possible that the subject as a whole is privy to the phenomenal characters of all of the states in JOINT, then would it not be possible that those phenomenal characters would all be experienced together? Would it not then be possible that all the states that compose JOINT do, in fact, have a conjoint phenomenology with each other?

It is extremely difficult to see these possibilities as coherent and it is hard to take them seriously. Nevertheless, for the theorist who is looking to understand the nature and structure of conscious experience, the notion of a spread subject cannot be dismissed merely because it is unorthodox or outside the bounds of our traditional conceptual framework. I believe that the above considerations on the basis

of subjective experience decide against the notion of a spread subject, but I admit that others may not find them to be as decisive as I do. Some two-stream views that will be evaluated below require that something like the notion of a spread subject is a coherent one. I will have more to say on the subject there. Here, I will merely reiterate my earlier point that the Consensus View does not by itself determine an intuitive or orthodox view about states of consciousness and their subjects. For that, we need a thesis about how many subjects a state of consciousness can have and we need an argument or reason to rule out the possibility of spread subjects of consciousness.

3.1.4 Streams of Consciousness

The intuitive notion of a maximal state of consciousness is conceptually related to the notion of a stream of consciousness. Having formalized the notion of a maximal state of consciousness, I will conclude this section by saying a few words about how maximal states relate to streams of consciousness.

A maximal state of consciousness is a state of consciousness that a subject has at a time. It consists of a collection of conscious states that are all experienced together with each other; every part of the maximal state has a conjoint phenomenology with every other part of the state. Moreover, maximal states are, in a sense, as big as they can be. A maximal state may not include every state that its subject has at the time, but it includes all of the subject's conscious states at the time that are experienced together as a package. A stream of consciousness is temporally extended. It unfolds over time. The intuitive notion of a maximal state of consciousness is that of a temporal slice of a stream of consciousness. In the rest of this essay, I will take the formal relation between a maximal state and a stream of consciousness to be nothing more than the intuitive one.

This view of the relationship between maximal states and streams of consciousness has at least one consequence: a stream of consciousness at a time, just like a maximal

state, need not include all of a subject's conscious states at that time. This just falls out of the Consensus View. Again, although this is inconsistent with our naïve view about the structure of consciousness, it is necessary that the formal machinery be set up in such a way that no questions are begged. Several of the views that will be discussed insist that subjects can have more than one stream of consciousness and this way of characterizing maximal states and streams of consciousness leaves this possibility open.

I have stipulated what the relationship between maximal states and streams of consciousness are, but this says very little, in any formal way, about what streams of consciousness are. The intuitive notion of a stream of consciousness is that of the sum of all of one's experiences over time. However, because the notion of a stream of consciousness, as it has been formalized so far, leaves open the possibility of a subject having more than one stream of consciousness, it departs significantly from the intuitive one. One might wonder if there is more that can be said about what a stream of consciousness is.

A stream of consciousness is a series of maximal states under a temporal ordering relation. But of course, not just any series of temporally ordered maximal states is a stream of consciousness. For starters, all of the maximal states that comprise a stream of consciousness are had by the same subject. In addition, streams of consciousness typically exhibit a certain diachronic unity relation. Speaking metaphorically, one might say that there is typically a flow to them; they come together to form a seamless temporal progression. This might give one the impression that the correct way to delineate streams of consciousness is in terms of maximal states of consciousness that are temporally sequenced, had by the same subject, and ordered by a kind of similarity relation; successive states in streams of consciousness are more similar to each other than states that are more temporally distant from each other. However, this way of characterizing streams of consciousness does not get things quite right.

Although streams of consciousness typically have a kind of flow and typically exhibit a kind of diachronic unity relation, there are cases in which streams of consciousness are quite disunified over time. It is even possible to imagine subjects of consciousness experiencing quite dramatic and sudden changes to the contents of their inner lives. In cases like this, it is quite possible for states that are temporally distant from each other to be more similar to each other than to states they are immediately temporally adjacent to. This would be less of a problem for the project of providing a precise characterization of streams of consciousness if it was not possible for a subject to have more than one stream of consciousness. However, once this is admitted to be a possibility, then individuating streams of consciousness in terms of the subjects of the states that compose them and the temporal relations those states bear to each other will not give the correct result.

Suppose a mad scientist gets a hold of a subject with two streams of consciousness and makes it so that the subject's two streams contain very different contents. Perhaps the mad scientist hooks up the subject's brain to the sensory systems of two separate individuals who are left free to carry on their lives in their customary manner. Presumably, the experiences of this two-stream subject would then be such that each of its two streams of consciousness would be diachronically unified over time even though there would be little or no overlap in terms of content. One can then imagine the mad scientist going further and deciding to make things so that the streams of information from the sensory systems of the two source individuals are randomly switched back and forth between the two streams of the two-stream subject. In a case such as this, ordering maximal states in terms of subject, time, and similarity would not produce the desired result, for it would fail to distinguish between a situation in which the mad scientist merely feeds the two-stream subject two independent streams of sensory information and the situation in which those streams of information are randomly switched back and forth.

The problem of characterizing streams of consciousness in a way that would get the right results in a case like this is a difficult one, but I will not take the time to attempt a solution here. My focus here, for the most part, will be on whether the notion of a two-stream subject is a coherent one. I will be arguing against it and my arguments will proceed, for the most part, by considering relations between maximal states of consciousness and their subjects. And if my arguments are successful, then the notion of a two-stream subject and the complications it brings will not need to be accounted for within a precise characterization of a stream of consciousness. I take it that maximal states have been characterized precisely enough for this discussion to proceed and so I will proceed on the assumption that the relation between a maximal state and a stream of consciousness is clear and that we have a clear enough sense of what a stream of consciousness is.

3.2 Split Brains and Two Streams of Consciousness

Having delineated a notion of a maximal state of consciousness that all parties in the discussion about the unity of consciousness can accept, and having said a few words about how a maximal state of consciousness relates to a stream of consciousness, it is now time to present, in a more precise way, and examine some of the different views about the nature and structure of the consciousness of split-brain subjects. The traditional and naïve view is that human subjects have just one stream of consciousness, and reflection on one's own experience may even lead one to suppose that consciousness is structured in this way for all creatures that enjoy conscious states. The data from split-brain subjects, however, seems to undermine this generalization. In the rest of this chapter, I will examine a series of views about the consciousness of split-brain subjects that are incompatible with the naïve and traditional view. These are all views according to which split-brain subjects have two streams of consciousness. I will argue that they should not be accepted.

3.2.1 Partial Unity

The first view I will take up is a view that might be characterized as a view according to which the consciousness of split-brain subjects is partially unified (or, equivalently, partially split). Perhaps the fullest defence and development of a partial-unity account is that of Lockwood (1989, 1994). Lockwood develops his partial-unity account around the idea that the co-consciousness relation is intransitive.

Lockwood asks his readers to consider what he calls a *phenomenal perspective* or a *maximal experience*. A phenomenal perspective is “an experience which does not form a part of a larger experience” (1989, 88). Although it is clear that Lockwood intends the notion of a phenomenal perspective to be equivalent to what I have called a maximal state of consciousness, this way of characterizing a phenomenal perspective is not equivalent to the way the Consensus View characterizes a maximal state of consciousness. The reason is that, unless one takes the composition relation to be restricted in one way or other, there is only one experience that is not part of another experience, namely, the sum of all experiences. Although the Consensus View assumes that experiences can be parts of other experiences, it does not have this entailment.

This difference between a phenomenal perspective and a maximal state of consciousness, however, seems to be only skin deep. The reason for this is that Lockwood seems to build considerations about conjoint phenomenology into his understanding of what an experience is. The notion of an experience, he says, “is to be understood in the philosopher’s slightly technical sense of a conscious state, happening, or sequence of happenings, that is experienced *as a whole*” (original emphasis) (1989, 87). Thus, although he does not explicitly address the question whether composition is restricted or not, he seems to be working with a notion according to which not just any collection of mental states composes a whole. As far as I can tell, the notion of a collection of experiences that is experienced as a whole and is not part of another collection

of experiences *is* equivalent to the notion of a maximal state of consciousness that is outlined by the Consensus View. For Lockwood, then, a phenomenal perspective is a mental state in which each part is *co-conscious* with each other part; a pair of conscious states are co-conscious when “there is an experience of which they are both parts” (1989, 88). Given that an experience is one whose parts are all experienced as a whole, a pair of co-conscious states will thus be a pair of states that have a conjoint phenomenology.

According to Lockwood, our naïve folk psychology is one according to which the co-consciousness relation is transitive; if a conscious state, c_1 , is co-conscious with c_2 and c_2 is co-conscious with c_3 , then c_1 must also be co-conscious with c_3 . Since the co-consciousness relation is also reflexive and symmetric, our naïve view is one according to which the co-consciousness relation neatly partitions conscious states into non-overlapping maximal states of consciousness (or phenomenal perspectives). Split-brain subjects, maintains Lockwood, show this view about the transitivity of co-consciousness to be mistaken; they show us that it is possible for a conscious state, c_1 , to be co-conscious with c_2 and c_2 to be co-conscious with c_3 *without* c_1 also being co-conscious with c_3 .

Data gathered from split-brain subjects reveals that the two hemispheres of the split-brain have a lot in common. For instance, although the contra-lateral hemisphere is often dominant with respect to motor control of the limbs, the ipsi-lateral hemisphere often has some motor control over the limbs. Both hemispheres are capable of controlling facial muscles. Both hemispheres receive sensory information from the neck, facial region, and from each ear. Both hemispheres perform relatively well on self- and social-knowledge tasks, and both hemispheres seem to respond emotionally in similar ways. However, as we saw in section 2.3.1, the data also suggest that there are experiences that may not be common to both hemispheres. If a pen is shown in the left visual field and a knife is shown in the right visual field, then, although the

experience of the pen seems to be conscious in the right hemisphere, it does not seem to be conscious in the left hemisphere, and vice versa with respect to the experience of the knife.

Lockwood's partial-unity view is an attempt to account for these facts. Suppose that, while being shown a pen in the left visual field and a knife in the right visual field, a split-brain subject is pricked on the neck. This would result in a situation, according to Lockwood, in which a visual image of the knife is co-conscious with a prick on the neck and the prick on the neck is co-conscious with a visual image of a the pen, but the visual image of the knife is not co-conscious with the visual image of the pen. The view is one according to which the consciousness of a split-brain subject consists of overlapping spheres of mutually co-conscious mental states. The visual image of the knife and the prick on the neck are part of one maximal state of consciousness, and the prick on the neck and the visual image of the pen are part of another maximal state of consciousness. Each of these two maximal states are, in turn, associated with distinct streams of consciousness. Because the maximal states partially overlap, the two streams of consciousness partially overlap as well. In this way, the consciousness of a split-brain subject is partially unified. However, because maximal states and streams of consciousness fail to overlap completely, the consciousness of a split-brain subject is also partially disunified.

Lockwood further motivates this partial-unity view by considering the following imaginary scenario. Suppose that, before the surgery, the commissurotomy subjects have but one stream of consciousness; their consciousness is (completely) unified. Suppose that after the surgery and the effects of the anaesthetic have worn off, they have separate, non-overlapping streams of consciousness, one associated with each brain hemisphere. Now imagine this surgery being performed using only local anaesthesia so that the subject of the surgery is conscious while the commissurotomy is performed. Lockwood asks what the consciousness of the subject would be like as

the individual neurons of the corpus callosum are severed one by one. At what point would the consciousness of the subject split into two? No specific answer to this question seems to be defensible. For any given answer, one could ask why the cutting of that specific neuron was responsible for the sudden emergence of the second stream of consciousness. Lockwood suggests that, given no particular answer to the question seems defensible, we should say, instead, that a disunified consciousness would emerge gradually. This would imply the possibility of a partially-unified consciousness.

Before moving on to consider some objections to Lockwood's partial-unity view, it is worth being absolutely clear on what the view is. A subject with a partially-unified consciousness has two maximal states of consciousness at a time and two streams of consciousness over time.⁷ Both the maximal states and their associated streams of consciousness partially overlap. They partially overlap by having parts that are not merely type-identical but *token*-identical.

3.2.1.1 Projecting Oneself into a Partially-Unified Consciousness

The partial-unity view seems to do a good job of accounting for the split-brain data. If split-brain subjects have two streams of consciousness, one for each hemisphere, then it would stand to reason that a split-brain subject might not be able to identify with one hemisphere a stimulus that is shown to the other; the mere fact that one hemisphere is shown a visual stimulus would not guarantee that a corresponding visual experience would be manifest in the stream of consciousness associated with the other hemisphere. The partial-unity view can also account for the fact that split-brain subjects get along so well outside of the experimental conditions. If a split-brain subject's two streams of consciousness overlap in the way that the partial-unity account maintains, then it would stand to reason that the two hemispheres could act

⁷A partially-unified consciousness could involve more than just two streams of consciousness, but since it is split-brain subjects that serve as the inspiration for the partial-unity view, the view is naturally presented as one according to which they have just two streams of consciousness.

in (largely) cohesive ways, for there would be a lot of content that would be similar across the two streams.

In spite of this, few are prepared to accept the partial-unity model. Dainton (2000, section 4.4), for instance, points out that it seems impossible to imagine what a partially-unified experience would be like. When we imagine a pair of co-conscious experiences we do so by attempting to hold the two experiences in mind simultaneously. Since a partially-unified consciousness is one in which the co-consciousness relation is not transitive, imagining a partially unified consciousness requires us to imagine c_1 together with c_2 and c_2 together with c_3 without, at the same time, imagining c_1 together with c_3 . But this seems impossible. Imagining c_1 together with c_2 and c_2 together with c_3 requires us to hold all three experiences in mind simultaneously, and this, by its very nature, is also an imagining of c_1 together with c_3 , a situation that constitutes an imagining of c_1 being *co-conscious* with c_3 .

Dainton cautions against ruling out a partially-unified consciousness on the basis of its being impossible to imagine. So does Bayne (2008). Bayne points out that many things that are unimaginable are, nevertheless, perfectly *conceivable*. For instance, because we do not navigate our world via echo-location in the way bats do, it may be impossible for us to successfully imagine what it is like to navigate the world in this way. Nevertheless, it is perfectly conceivable that a creature could navigate the world via echo-location. Bayne goes on to argue, however, that not only is partial unity unimaginable, it is also inconceivable. He says, “Although the phenomenal perspective of a bat is not imaginatively accessible to us, we have no difficulty conceiving that there is something it is like to be a bat. By contrast, we have difficulty—arguably *great* difficulty—in conceiving that there is something it is like to be partially unified” (original emphasis) (2008, 293).⁸

⁸It is worth noting that Lockwood (1994, 95), writing several years after first presenting the partial-unity account, expressed his own doubts about its coherence.

Both Dainton and Bayne make some comments suggesting that the reason partially-unified consciousness is inconceivable has to do with the way in which we are acquainted with the relation of co-consciousness. Dainton points out that, whatever properties the relation of co-consciousness has, we “know *what it is like* for experiences to be so related” (original emphasis) (2000, 105). It is because we know what it is like for experiences to be related via the co-consciousness relation that we are inclined to reject the idea that co-consciousness is intransitive; our experiential knowledge of the relation makes it so that partial unity strikes us as incoherent. Bayne says something very similar to this. He says that “[a]rguably, first-person acquaintance with consciousness reveals that simultaneous phenomenal states that are unified with a third (simultaneous) experience *must* be unified with each other” (original emphasis) (2008, 293). These observations strike me as sound. Our acquaintance with the relation in our own experiences is what anchors our understanding of what its formal properties are. Our experience seems to tell us that co-consciousness is transitive, and so a situation in which co-consciousness is intransitive is inconceivable. If partial unity is, in addition to being unimaginable, inconceivable, there is reason to reject the partial-unity model.

One problem with this kind of argument for the inconceivability of an intransitive co-consciousness relation is that, given our knowledge and understanding of the relation is so rooted in experience, it is difficult to say much about it that will convince someone who is not already in agreement. Indeed, not all share the inconceivability judgement. Tye (2003), for instance, suggests that there is nothing particularly problematic about co-consciousness being intransitive. Consider again the split-brain subject that is shown a pen in one half of the visual field and a knife in the other while being pricked on the neck. According to Tye, there is no great mystery here; the split-brain subject has “two multi-modal experiences” (2003, 130). The subject has one experience that consciously represents the pen together with the prick and an-

other experience that consciously represents the knife together with the prick. What the subject lacks is an experience that consciously represents the pen together with the knife.

There is thus no more mystery attaching to the non-transitivity of phenomenal unity than there is to the non-transitivity of the relation of entering into the same thought content. Just as I can think that Ann loves Paul and also think that Paul is the neighbour of Raoul without having any thought into whose content Ann and Raoul enter, so a split-brain patient can experience *A* unified with *B* and also experience *B* unified with *C* without his experiencing *A* and *C* together. (2003, 131)

I suggest that the reason Tye finds nothing particularly problematic about the non-transitivity of phenomenal unity (or co-consciousness) is that his account of the consciousness of the split-brain subject is not necessarily one according to which co-consciousness is intransitive. The account Tye gives of the split-brain subject is consistent with a situation in which the subject has two maximal states that are merely qualitatively similar, not necessarily one in which the two maximal states partially overlap by sharing token-identical states. There is nothing inconceivable, problematic, or mysterious about two distinct maximal states being partially type-identical. If two people stand beside each other in an art gallery and look at the same spot on a painting on the wall, their experiences will have a large degree of qualitative similarity. But nothing about this says anything about the transitivity, or lack thereof, of co-consciousness or phenomenal unity. We do not believe that because the visual experiences of the two people have a lot in common that they share (token-identical) mental states. What partially-unified consciousness requires is a case in which the very same experiential particular is co-conscious with each of two other experiential particulars that are, in turn, not co-conscious with each other. As long as we are considering cases that involve two non-overlapping maximal

states, we are not considering a case that is of the right sort to exemplify a partially-unified consciousness. Two people looking at the same expanse of red on a painting both have a conscious representation of that expanse. However, their respective conscious representations of that red expanse are numerically distinct experiential particulars. Thus, even if the one person had an experience of red together with an experience of a sore left foot and the other person had an experience of red together with an experience of a sore right foot, we would not have a case of partially-unified consciousness. The one person's experience of red would be an experiential particular that was numerically distinct from the other person's experience of red, even if the two experiences of red were qualitatively identical. This is why, as soon as we describe the case of the split-brain subject in the way Tye seems to do, we do not have a case in which co-consciousness or phenomenal unity is intransitive. For all that has been said, Tye's case is one in which a subject has two streams of consciousness, both of which exhibit *transitive* co-consciousness.

It is difficult, perhaps even impossible, to imaginatively project oneself into a partially-unified consciousness, and it is also very difficult, *pace* Tye, to conceive of a partially-unified consciousness. Yet imaginability and conceivability issues are often not easily decided. Imaginability and conceivability are, ultimately, somewhat dependent upon the individual doing the imagining or the conceiving; what may seem to one to be imaginable may seem to another to be unimaginable. If the defender of the partial-unity view claims, with apparent sincerity, that he or she finds a partially-unified consciousness to be conceivable, then we may have arrived at a stalemate.

I have argued that Tye's account of why there is nothing particularly inconceivable about a partially-unified consciousness fails to show what Tye claims it does. Although I will not have more to say by way of direct argument against the partial-unity view, some of the arguments I will make below against other interpretations of the split-brain data will also tell against the partial-unity view. The interpretations

of the split-brain data I will consider below are like the partial-unity account in that they claim split-brain subjects have two streams of consciousness. They differ from the partial-unity account in their assertion that the two streams of consciousness do not overlap. This makes them immune to some of the conceivability worries that plague the partial-unity account. However, as I will argue below, at least one way of specifying this kind of alternative to the partial-unity view may still face some challenges on this front. The challenge is to say what the substantive difference is between a single subject having two streams of consciousness (overlapping or not) as opposed to just one. I will argue in the next section that there is no adequate response to this challenge and, thus, that the conceivability of a single subject having two streams of consciousness is called into question. This will tell against the partial-unity account, because it, like some accounts that posit non-overlapping streams of consciousness, relies crucially on the notion that it is possible for a single subject of experience to have two distinct streams of consciousness.

3.2.2 Two (Disjoint) Streams

The partial-unity account claims that split-brain subjects have two streams of consciousness that overlap. In virtue of this overlap, the two streams will have content in common, but the partial-unity after which the view has its name is a unity due to the literal sharing of mental states across streams of consciousness. An alternative to the partial-unity account is to claim that, although the partial-unity account is correct in attributing two streams of consciousness to split-brain subjects, it is incorrect in its assertion that the streams overlap. According to this view, the consciousness of split-brain subjects is completely disunified. This is not to say that the two streams of consciousness of split-brain subjects may not enjoy a large degree of overlap in a different sense—they may well have a large amount of *content* in common—however, no conscious state or experience in one stream is co-conscious with any conscious

state in the other stream. According to this view, each conscious state is co-conscious with every other state *within* a stream of consciousness, but all states related to each other by the co-consciousness relation are perfectly partitioned into non-overlapping streams; co-consciousness is transitive. I will refer to accounts of this sort as “two-stream” accounts. This way of referring to them is somewhat misleading in that it implies that one of their main rivals, i.e., the partial-unity account, does not attribute two streams of consciousness to split-brain subjects. As we have seen, however, the partial-unity account does attribute two streams of consciousness to split-brain subjects. For this reason, it might be helpful to think of the label, “two-stream,” as carrying a tacit qualifier to the effect that the two streams are completely disjoint.

Two-stream proposals, like the partial-unity proposal, are motivated by the apparent differences in conscious content that seem to be generated by the two hemispheres of split-brain subjects in the experimental conditions. If one hemisphere seems to be unable to identify an object that is presented to the other, then there is reason to suppose that some conscious content that is generated by one hemisphere is not generated by the other. Two-stream proposals do at least as good a job, if not a better one, at accommodating this consideration than the partial-unity proposal does. Thus, there is at least one reason to take them seriously. In addition, two-stream proposals avoid some of the conceivability worries that bedevil the partial-unity account; two-stream accounts need not deny the transitivity of co-consciousness and so do not need to account for how a pair of overlapping maximal states could fail to count as a single maximal state.

Two-stream proposals, however, come with their own unique difficulties. In non-experimental circumstances, split-brain subjects appear perfectly normal. They do not seem to suffer from any perceptual difficulties and their behaviour is at least as coherent and unified of purpose as that of others. This is somewhat surprising given that, for some of the senses, most notably vision, there is a significant disparity be-

tween the sets of information that the right and left hemisphere receive from the right and left visual fields. If split-brain subjects have two distinct streams of consciousness that may diverge significantly in terms of content, what explains the coherence and unity of behaviour that split-brain subjects exhibit outside of the experimental conditions? It has been demonstrated that, although there is often an asymmetry between how much motor control the respective hemispheres have upon the limbs—with the hemispheres being more dominant with respect to motor control of the contra-lateral limbs—both hemispheres are able to exert at least some degree of motor control upon the limbs on both sides of the body. If, as is quite likely, the brain hemispheres of split-brain subjects receive differing sets of sensory information, one might well expect there to be more conflict in terms of how the limbs are manipulated.

I have been speaking about two-stream accounts without having said anything about how two-stream accounts might differ from each other. There are different ways in which two-stream accounts might be further specified and it will be worth distinguishing these ways now. Defenders of these different two-stream accounts are likely to respond to the above question in different ways and so it will be necessary to consider their views separately. Two-stream accounts are unanimous in claiming that split-brain individuals have two disjoint streams of consciousness. Where two-stream accounts may part ways, however, is over the number of subjects of experience split-brain individuals house. According to one specification of the two-stream proposal, there are two streams of consciousness but only one subject of experience. According to another way of specifying the two-stream account, however, there are two subjects of experience, one for each stream of consciousness. I will refer to these different two-stream accounts as the TSOS (two streams, one subject) and TSTS (two streams,

two subjects) accounts respectively.⁹ I will consider the TSOS account first and then consider the TSTS account.

3.2.2.1 Two (Disjoint) Streams and One Subject

The TSOS account appears to be in a good position to account for the behavioural unity of split-brain subjects outside of the experimental conditions while at the same time accommodating the fact that the two hemispheres have different sets of sensory information available to them. The defender of the TSOS account can explain the behavioural unity of split-brain subjects outside of experimental conditions with the fact that the two streams of consciousness have the same subject. Although the two streams of consciousness do not overlap in terms of their content, the two sets of content are both available to the same subject of experience. The (single) subject of experience can then take all of the content that is available to it via the two streams and produce singular intentions and coordinate a unified set of motor commands.

There are several problems with the TSOS account. The first has to do with the fact that accounting for the behavioural unity of split-brain individuals outside of the experimental conditions in this way causes trouble for the view when the behavioural *disunity* within the experimental conditions must be accounted for. In order to account for behavioural unity outside of the experimental conditions, the locus of behavioural intentions and motor commands associated with the subject of experience; there is one subject of experience and, thus, one set of behavioural outputs. If the locus of behaviour is associated with the subject of experience in this way, then it is puzzling why split-brain subjects should exhibit such strikingly disunified behaviour within the experimental conditions. If the (single) subject of

⁹Defenders of two-stream accounts seldom differentiate between TSOS and TSTS accounts and often seem to slide back and forth between them, leaving it unclear what view they take themselves to be defending. This is problematic because, as will become clear in the following pages, these two ways of understanding the two-stream account have different implications and are subject to different objections.

experience has access to the content from both streams of consciousness when outside of the experimental conditions, why not also when in the experimental conditions? Partitioning experience into two separate streams has the benefit of producing a structure that seems to do a nice job of accounting for the differing sets of sensory information that is processed in the two hemispheres, but if the content of these two streams of consciousness is available to the same subject of experience and, moreover, the subject of experience is made the locus of behavioural outputs, the experimental data becomes much more difficult to account for in a consistent way.

A second problem with this view, as Bayne (2008) points out, is that it seems to be incompatible with a relatively standard view about how conscious states are individuated. A standard way to individuate conscious states is in terms of their content, subject, and time. Distinct conscious states can have any two of these in common, but states that have all three in common are identical. The TSOS account is motivated, in large part, by considerations about what sensory information seems to be available to the respective hemispheres. Although there are clear differences between, say, the visual information that is processed in the respective hemispheres, there is also, as has been pointed out, a non-trivial amount of overlap between the sensory information available to the two hemispheres. If the physical basis for the two streams is taken to be the two hemispheres respectively and the content of the two streams is taken to be in one-to-one correspondence with the sensory information processed in the respective hemispheres, then the defender of the TSOS account must say that the two streams of consciousness overlap somewhat in terms of content. There is nothing incoherent or problematic about distinct streams of consciousness having content in common. If we consider, again, the situation in which two people are standing beside each other in an art gallery looking at the same point on a painting on the wall, we have every reason to believe that their respective streams of consciousness will be very similar in terms of qualitative character and overall content. However, the

defender of the TSOS account maintains that the distinct streams of consciousness have the same subject. This is where the TSOS account runs into trouble with the standard way of individuating conscious states. In a situation like the one described by Tye, in which a split-brain subject is shown a pen and a knife in opposite visual half-fields while being pricked in the neck, the defender of the TSOS account must say that the prick of the neck is manifest in both streams of consciousness. But the sensation of the prick in the one stream has the same subject as the sensation of the prick in the other stream. This has the result that, on the standard tripartite account of how conscious states should be individuated, the sensations of the prick are not two conscious states but one; they are identical in terms of their qualitative character, they occur simultaneously, and they have the same subject. The standard tripartite account of how conscious states should be individuated leaves no room to make sense of a single subject having two qualitatively overlapping streams of consciousness. In this one regard, the partial-unity account has a leg up on the TSOS account; the partial-unity account claims that the two streams have content in common, but that is because the conscious states themselves are shared by the two streams. Of course, as we have seen, the partial-unity account has other problems, but it is consistent with the tripartite account of how conscious states should be individuated.

In my view, there is a third and even deeper problem with the claim that a single subject of experience could be the subject of two numerically-distinct streams of consciousness. The problem is similar to the problem that arose for the partial-unity account when it came to conceiving of a scenario in which the sum of all of a subject's conscious states at a time could fail to be a maximal experience. According to the TSOS account, some of split-brain individual's experiences belong to one conscious stream and the rest to belong to another conscious stream. Yet the simultaneous experiences in the two streams of consciousness would seem to satisfy all of the conditions for co-consciousness: they occur simultaneously and have the same subject.

What is there that these experiences lack such that, if they had it, they would be co-conscious? It is not easy to see what that might be. And if they are all, in fact, mutually co-conscious, then it is not clear what could be the basis for maintaining that there are two, rather than one, conscious streams.

In my discussion of the partial-unity account, I considered Tye's explanation for why there is nothing particularly inconceivable about a partially-unified consciousness. The verdict was that Tye failed to show us that a partially-unified consciousness is conceivable. I will do a similar thing here with respect to the TSOS account. I maintain that the TSOS account fails the conceivability test just like the partial-unity account does, though perhaps not for the same reasons. Before I move on to consider two-stream accounts that posit two subjects of experience, I want to say something about why one might be inclined to believe that it is conceivable for a single subject of experience to have two distinct streams of consciousness and then say a few things about why this belief is not well-founded.

3.2.2.2 Being Led Astray by Spatial Unity

Consider a scenario involving two individuals. Setting aside, for the moment, questions about imaginability and conceivability, let us stipulate that the first has two completely disjoint streams of consciousness. The second has only one stream of consciousness. Both subjects are given qualitatively identical sets of sensory information. We will suppose, to keep matters simple, that they have simultaneously been shown a pen in the left visual field, a knife in the right visual field, and received a prick on the neck. The question I want to ask about these two subjects is what the substantive difference between their respective inner lives is at that moment.

In the literature, maximal states of consciousness are often pictorially or graphically represented with circles. A partially-unified consciousness is represented by drawing a pair of circles that partially overlap (Figure 3.1). What the circles rep-

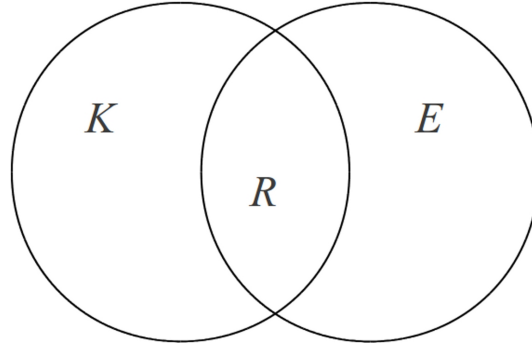


Figure 3.1. Two Overlapping Streams

resent is a boundary that includes all of the members of a set of conscious states that are mutually co-conscious with each other; all the states within the circle have a conjoint phenomenology. The region of overlap represents a set of states that inhabit both sets of mutually co-conscious mental states. A partially-unified consciousness can then be contrasted with a situation in which two streams of consciousness are completely disjoint (Figure 3.2) and a situation in which there is only one stream of consciousness (Figure 3.3). The consciousness of the two subjects in the scenario we

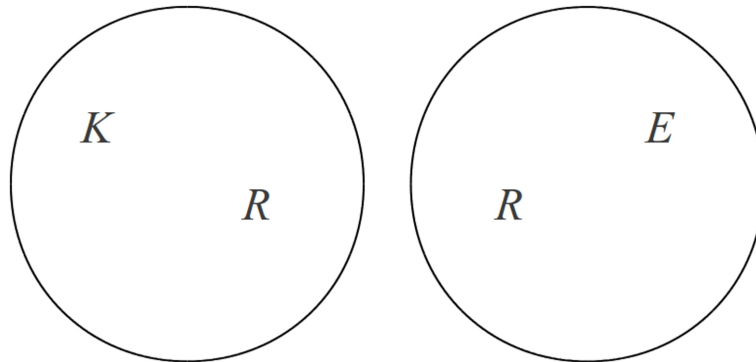


Figure 3.2. Two Non-Overlapping Streams

are considering are thus represented by Figure 3.2 and Figure 3.3 respectively. Both subjects have a visual experience of a pen, E , a visual experience of a knife, K , and an experience of a prick on the neck, R . The fact that the circles representing the two maximal states of consciousness of the first subject are non-overlapping is intended to

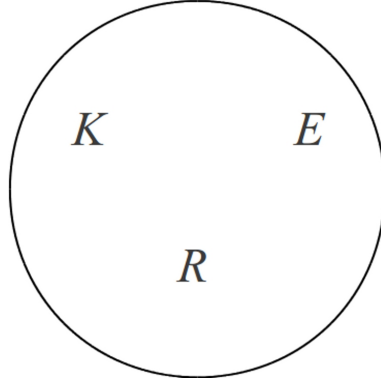


Figure 3.3. One Stream

represent a kind of sharp phenomenal divide between the two maximal states. This divide is supposed to represent the crucial and essential difference between being a subject with two streams of consciousness and a subject with only one. The question is what does this sharp divide amount to? What difference does it make from the inside?

A quick answer would be to repeat one of the talking points of the debate: it marks the fact that the subject with two streams of consciousness has two sets of mental states that are not co-conscious with each other. But it is hard to see how this answer points to a substantive difference in the inner lives of the two subjects. The reason is similar to those discussed above regarding the imaginability or conceivability of a partially-unified consciousness. The issue there was that, if a pair of states, c_1 and c_2 , are co-conscious and c_2 is also co-conscious with c_3 , then it seems that c_1 would also have to be co-conscious or experienced together with c_3 . The situation is not exactly the same with respect to non-overlapping streams of consciousness, for they do not have any mental states in common, but similar considerations apply. If c_1 is co-conscious with c_2 , and both c_1 and c_2 are experienced by the same subject, s , and c_3 is co-conscious with c_4 , and they too are experienced by s , it is hard to see how the states in these two co-conscious pairs could fail to be experienced *together* by s . What could it mean for a single subject to have a number of experiences

without those experiences all being experienced together as part of a single state in which their respective phenomenal characters come together to form a state with a total phenomenal character that includes them all? No doubt there will be those who claim that it is possible to see what the substantive difference between the two subjects in the scenario I have described is. It is my suspicion that those who are of this opinion are being led astray by a way of thinking about what it is like for the two-stream subject in this scenario that is irrelevant with respect to phenomenal unity.

It is easy to think of the divide (in Figure 3.2) in spatial terms. The representation is pictorial and represents what it is intended to represent via spatial relations. Thus, when thinking of the representation, it is easy to think of the circles as representing spatial boundaries; it is easy to think that the circles are spatial “containers,” indicating that the points within one circle are spatially related to each other but that they are not spatially related to the points in the other circle. And when the pictorial representations are thought of in this way, it is easy to conceive of an experience that is consistent with what the diagram is understood to represent.

Our experience is typically spatial; the objects we perceive are typically experienced as being part of a three-dimensional realm. Kant thought that experience is necessarily spatial in this way. Many question whether Kant was right about this, but it is certainly true that much of our experience is, as a matter of *actual* fact, experienced in this way. One way to try to imagine what it would be like to be a subject with two disjoint streams of consciousness is to imagine having two sets of experiences in which the objects perceived via those experiences are not perceived as being spatially related to each other.

There is a large degree of overlap between the visual fields of our two eyes. If we close one eye and take note of what we see with the other and then switch so that the open eye becomes closed and the closed eye becomes open, our visual experience

remains largely constant through the switch. Depending on which eye happens to open, the visual image will extend further to one side than the other, but there remains a large degree of overlap in the middle. One can imagine, however, having eyes situated much further apart, perhaps like those of some birds and reptiles whose eyes are situated on the sides of their heads. These animals have a much smaller degree of overlap between the visual fields of their two eyes. It is but a small further step to imagine having eyes situated in such a way that their respective visual fields do not overlap at all. However, merely having eyes with non-overlapping visual fields would not by itself ensure having the kind of experience we are attempting to imagine. It would still be possible to perceive the objects seen in the non-overlapping fields as being situated in the same physical space and as being spatially related to each other. For instance, the mental mechanism responsible for placing the experienced items in three-dimensional phenomenal space¹⁰ could take the non-overlapping nature of the two visual fields into account and situate the experienced items in the same phenomenal space even though the objects in the two visual fields would never be seen by both eyes simultaneously.¹¹ But once we see how a mental mechanism could compensate for non-overlapping visual fields, it is but a small further step to see how a creature with non-overlapping visual fields could fail to perceive the objects seen in the two visual fields as being located in the same space.

For creatures like us who are accustomed to seeing objects that are spatially related to each other it may be a bit of a challenge to imaginatively project ourselves into such an experiential reality, but it is certainly not inconceivable. Perhaps it is more easy to imaginatively project oneself into a situation in which the spatial disconnect

¹⁰Dainton (2000) distinguishes between phenomenal space and physical space. Phenomenal space is generated by the mind and can correspond, to greater or lesser extents, to physical space. I am attempted to follow Dainton's usage here.

¹¹If it is difficult to see how this could be possible due to the fact that a lack of overlap would seem to compromise depth perception, imagine having two *pairs* of eyes situated in such a way that the field of view for one pair of eyes does not overlap with the field of view for the other pair.

occurs between two sets of experiences that come from different sensory modalities. One can imagine, for instance, having one's optic nerves severed and connected via sophisticated radio transmitters to visual input devices at some remote location so that one's visual experience in no way lines up with what one experiences via one's other senses.¹² The objects seen in such a scenario could well be experienced as occupying a spatial realm that is distinct from the objects perceived via one's other senses. But this kind of spatial disconnect between perceived objects need not merely be drawn from the realm of imagination. One can actually experience it. For instance, it is often difficult to determine, in a large city while surrounded by skyscrapers, where an emergency vehicle is coming from. Sometimes the siren is heard as coming from one direction when, in fact, it is coming from another. However, I often find that I can not even say for certain, one way or another, which direction the siren *seems* to be coming from. In that moment, it is not possible for me to tell, merely on the basis of my experience, where in phenomenal (or even physical) space the emergency vehicle is situated relative to the objects I perceive visually in front of me.

It is easy to think that the inner life of a subject with two streams of consciousness is like this. That is, it is easy to think that having two disjoint streams of consciousness would be a little like having two sets of experiences in which the objects perceived in them respectively are not perceived as being spatially located to the objects in the other experience. When thought of in this way, it is easy to see how the inner life of a subject represented by Figure 3.1 could be different from a subject represented by Figure 3.2; the first subject in the scenario I described earlier would not experience the pen as being spatially related to the knife. The subject would probably also experience the pricks in the two streams as being numerically distinct. The experience of the

¹²For some similar science fiction examples, see Dainton (2000) and Dennett (1978). What is interesting is that these kinds of cases are no longer purely the stuff of science fiction. See, for instance, the work of O'Doherty et al. (2011) and Pais-Vieira et al. (2013).

second subject would be very different in character, because the second subject would experience the pen and the knife as being spatially related to each other.

This way of thinking about the putative difference between the experiential lives of a subject with two streams of consciousness and a subject with only one is mistaken; it does not point to a genuine difference between what it would be like to have two streams of consciousness as opposed to just one. The reason for this is that it is unclear that a subject would have to have two streams of consciousness in order to experience objects as being spatially isolated from each other. Once we see that it might well be up to a mental mechanism to ensure that objects are perceived as being related to each other within a three-dimensional space, there seems to be no particular reason to suppose that such a mental mechanism could be absent or deficient in a subject with only one stream of consciousness. For instance, there does not seem to be anything particularly inconceivable or even unimaginable about an otherwise normal human perceiver having a condition whereby the objects perceived in one region of the visual field are not perceived as being spatially related to the rest of the objects perceived via the visual system. There is nothing about this condition in and of itself that should lead us to suppose that such a subject would have two non-overlapping streams of consciousness. If it is not necessary to be the subject of two streams of consciousness in order to live within this kind of perceptual reality, then all this business about experiencing objects as spatially isolated does not point to a substantive difference between having one stream of consciousness or two.

There is another feature about the case we have been considering that the defender of the TSOS account might attempt to use to point to the kind of difference that having two streams of consciousness might make. Consider again Figure 3.2. “*R*” occurs in both spheres. This represents a situation in which there are two numerically distinct states, one in each sphere or maximal state, that are type-identical. What would it be like for a subject to simultaneously have two experiences that are type-

identical? Surely this would amount to a kind of double experience. Would not the subject experience two pin pricks rather than one?

Pointing to this feature of the cases as an example of the kind of experiential difference that having two streams of consciousness could make, however, would be premature. The reason, again, is that one does not have to be the subject of two streams of consciousness to enjoy a pair of type-identical experiences. We typically experience pricks as being located at a particular place on the surface of our body, and therefore it might seem difficult to imagine having two qualitatively identical experiences of pin pricks. However, there are many places on the surface of our bodies where our sensitivities to things like pin pricks are relatively course-grained. All one needs to do is imagine being pricked in such a way that the pricks are experiences as being two while, at the same time, it being difficult to distinguish their location relative to each other. It is also quite easy to come up with other situations in which one can have simultaneous type-identical experiences. One can look at an expanse on a painted wall that is evenly lit and have experiences that are type-identical—the colour experience on the left may be type-identical to the colour experience on the right. One can hold an ice cube in each hand and also have indistinguishable experiences. Although the number of experienced pin pricks may point to a difference between the experiential lives of the two subjects in the scenario we considered above, it does not point to a difference that is constitutive of having two streams of consciousness as opposed to one.

3.2.2.3 Subjects and Conjoint Phenomenology

I believe these considerations show more than just that spatial disunity should not be confused with phenomenal unity or that double-experience is possible within a single stream of consciousness. Once we begin to examine how one might come to the belief that partial-unity or TSOS accounts describe a legitimate possibility

for the structure of consciousness and come to see why these considerations do not necessarily decide in favour of these accounts, it becomes more and more plausible that being experienced simultaneously by the same subject may be sufficient for a pair of mental states to be co-conscious. And once we begin to see how hard it is to say what more is needed for a pair of conscious states to have a conjoint phenomenology than for them to be simultaneous and have the same subject, we begin to see that the following may well be true.

For any pair of (simultaneously occurring) conscious mental states, s_1 and s_2 , if s_1 and s_2 have the same subject, then s_1 and s_2 have a conjoint phenomenology. (3.6)

Although (3.6) is not accepted by defenders of partial-unity and TSOS accounts, all, aside from those who question the very existence of subjects of experience, accept the claim that having a conjoint phenomenology entails sameness of subject. Thus, to the extent that (3.6) is plausible, the following bi-conditional version of (3.6) is also plausible.

For any pair of (simultaneously occurring) conscious mental states, s_1 and s_2 , s_1 and s_2 have the same subject if and only if s_1 and s_2 have a conjoint phenomenology. (3.7)

An interesting consequence of (3.7) is that, if it is true, maximal states could be defined either in terms of conjoint phenomenology or in terms of sameness of subject. Another way to say this is that, if (3.7) is true, one could define the notion of a maximal state either by omitting clauses (i) and (ii) of the Consensus View or by omitting clause (iii).

I maintain that the absence of a plausible account of what might be the subjective or internal difference between a state of affairs consisting of a single subject with only

one stream of consciousness and a state of affairs consisting of a single subject with two streams of consciousness (overlapping or not) gives us reason to take (3.7) seriously. Indeed, I am inclined to accept (3.7) outright. Those who defend partial-unity and TSOS accounts, however, may greet the considerations I have outlined in the previous sections with a shrug, and there is one response they might give to them that is worth saying a few words about here.

I have been examining different ways one might attempt to make sense of what the structure of a disunified consciousness of the sort suggested by partial-unity and TSOS accounts is. I have argued that whatever might be offered as a putative difference between a subject having one stream of consciousness and a subject having two is not a genuine difference; features that may seem to be unique to a structure of consciousness involving a single subject with two streams of consciousness turn out to be possible within a single stream of consciousness. Defenders of partial-unity and TSOS accounts, however, might argue that I have been begging the question against them. They might argue that I have smuggled the notion of conjoint phenomenology into my notion of a subject of consciousness and that, in my efforts to evaluate partial-unity and TSOS accounts, I have conveniently built more into the notion of a subject of experience than the nature of this debate allows. If I have assumed, at the outset, that all of a subject's experiences at a time must be related via the relation of co-consciousness, then of course it will turn out that theses like (3.6) and (3.7) will appear to be true. This is a significant charge. Should it turn out to be well-founded, my case against partial-unity and TSOS accounts will be significantly weakened.

I do not believe that I have begged this question. The Unity Thesis is, first and foremost, a thesis about whether all of a subject's conscious states at a time compose a single maximal state of consciousness. But maximal states are defined in terms of conjoint phenomenology, and so the Unity Thesis is ultimately a thesis about certain

phenomenal facts.¹³ Because co-consciousness (i.e., having a conjoint phenomenology) is a phenomenal relation, one can justifiably expect that rival views about how the relation is instantiated in split-brain individuals should entail substantive phenomenal differences. Thus, questions about what it would be like to be the subject of two streams of consciousness are fair questions. If, as I have argued, the answers to these questions make it difficult to see what the phenomenal difference is between having two streams of consciousness or having only one, then we have good reason to wonder whether partial-unity or TSOS accounts describe genuine alternative structures of consciousness.

Moreover, these questions can be asked without presupposing more than the bare minimum with respect to subjects of consciousness. The notion of a subject of consciousness is a mysterious one. Although we all have an intuitive sense of what it means to be the subject of a conscious state, the notion of a subject of experience is one that resists characterization in more basic terms. This is significant in the context of this discussion because, if it resists characterization in more basic terms, we have reason to believe that it might be a conceptual primitive. And if it is a conceptual primitive then the notion of conjoint phenomenology cannot be part of its conceptual constitution. Of course, people often employ subtly different notions within their philosophical analyses, and the question, here, is whether I have used such a notion. I do not believe I have. In order to attempt to understand what it would be like to be the subject of two streams of consciousness I considered different ways in which the idea could be fleshed out, and this involved considering situations that involved a subject having a number of different conscious states. The mere having of a conscious state does not, at least not in any obvious way, entail anything about whether

¹³Clause (ii) of the Consensus View mentions subjects of consciousness, but it says nothing about whether or not a subject can have more than one maximal state at a time. That claim is made by the Unity Thesis.

or not that conscious state is accompanied by any other conscious states or anything about the phenomenal relations that may or may not obtain between them. Thus, merely asking what further conditions, other than being had by the same subject simultaneously, a pair of conscious states need to satisfy in order to have a conjoint phenomenology does not assume that whatever conscious states a subject has must all have a conjoint phenomenology. The charge that I have begged the question is not well-founded.

I believe that, absent a plausible account of what might be the subjective or internal difference between a state of affairs consisting of a single subject with only one stream of consciousness and a state of affairs consisting of a single subject with two streams of consciousness (overlapping or not), we have reason to take (3.7) seriously. And if the absence of such an account does not give us enough reason to accept (3.7) outright, we must ask whether there are other reasons for not doing so. Defenders of partial-unity and TSOS accounts will, of course, point to the experimental findings from split-brain subjects. But given the fact that there are other interpretations of the data on offer, that is, interpretations that do not require one to posit the possibility of a single subject of experience having two streams of consciousness, we should probably look to them before rejecting such a compelling thesis as (3.7).

3.2.2.4 Spread Subjects Revisited

The TSOS account also requires that the notion of a spread subject is a coherent one. In section 3.1.3, the notion of a spread subject came up in a discussion of some non-standard views about the structure of consciousness that the Consensus View does not rule out by itself. The question was asked whether the Consensus View rendered a verdict on *JOINT*, i.e., the sum of my current conscious states and the sum of your current conscious states. The answer was that, if the notion of a spread subject was coherent, then the Consensus View does not entail that *JOINT* is a maximal state.

The TSOS account requires that a single subject can be the subject of two streams of consciousness simultaneously. Thus, the TSOS account invokes spread subjects as part of its account of experiential lives of split-brain subjects.

This presents a challenge to the defender of the TSOS account. If we are to believe that a split-brain individual is the subject of two non-overlapping streams of consciousness, must we also admit that there might be a single spread subject of JOINT? Although the TSOS account posits spread subjects, I find it doubtful that defenders of the account will want to accept the existence of arbitrary spread subjects that span two or more bodies; it is one thing to suggest that a spread subject might span the two hemispheres of a single brain, but it is quite another to suggest that there might be a single subject of something like JOINT.

There is a ready response for the defender of the TSOS account. In the case of split-brain individuals there is a much more unified set of behaviours and intentional actions. In the case of the putative subject of JOINT, there is no discernible unified set of behaviour; it does not appear as though there is a single subject directing the actions of the bodies responsible for its conscious experiences and over which it has intentional control. This is the obvious line to take for the defender of the TSOS account, but it comes at the cost of making the account somewhat unstable. Split-brain individuals demonstrate behavioural disunity in the experimental conditions. If behavioural disunity is the rationale for saying that JOINT is purely a philosopher's creation and consists, ultimately, of two maximal states that each have their own subject, then the defender of the TSOS account is under some pressure to abandon it and take up the view that I will be discussing next: the TSTS account.

3.2.2.5 Two (Disjoint) Streams and Two Subjects

Two-stream views are motivated by the behavioural disunity the split-brain individuals demonstrate within the experimental conditions. But these individuals also

demonstrate a large degree of behavioural unity outside of the experimental conditions. The TSOS account is motivated in large part by this behavioural unity. Nevertheless, many defenders of two-stream accounts claim that split-brain individuals house two subjects of experience, one for each stream of consciousness. In this section, I will take up this view.

The TSTS account, perhaps more so than any of the other accounts we have considered to this point, clashes with our naïve psychological views. For instance, it challenges naïve ideas about how many subjects of experience are associated with a single body (or brain). Although defining precisely what constitutes personhood remains an elusive goal, few doubt that the notion of personhood and the notion of a subject of experience are tightly bound. Maintaining that there are two subjects of experience housed in a split-brain subject's body seems to set one squarely down the path of having to seriously entertain the view that split-brain subjects are, in fact, two people.

What makes matters even more difficult is that once it is said that there are two subjects of experience in commissurotomy patients, it becomes difficult to resist the implication that there are two subjects of experience in normal brains with intact corpus callosums. A commissurotomy leaves the neural structures in the respective hemispheres largely intact. If a hemisphere is capable of sustaining a stream of consciousness after the severing of the corpus callosum, there is little reason to suppose that it was not capable of doing so beforehand. The only thing the commissurotomy seems to do is bring about a change in what information is presented and processed in each hemisphere. It would not seem to affect the capacities of the hemispheres to generate streams of consciousness. If each of the hemispheres generates a stream of consciousness after the commissurotomy, why not before? Some are prepared to bite this bullet (see Puccetti, 1981), but it does come with a significant theoretical cost.

Perhaps the most pressing concern for defenders of TSTS accounts is to say something about how individuals that house two subjects of experience could demonstrate such behavioural unity in their day-to-day lives. The most promising response to this worry is to maintain that there is a large amount of content duplication across the two streams. How else would the two subjects be able to act in such unified and complimentary ways? There is, as has been mentioned several times already, nothing incoherent about two streams of consciousness overlapping significantly in terms of content.

Is this plausible? There is reason to be sceptical. In normal individuals, the corpus callosum serves to courier much of the sensory information between the two hemispheres that would otherwise be restricted to one hemisphere or the other. Split-brain individuals do not have this information pathway. This does not mean, however, that the two hemispheres of split-brain individuals do not receive *any* sensory information from the ipsi-lateral sensory field; there is, for instance, evidence that the two hemispheres receive visual information from the peripheries of both visual fields.¹⁴ Although there is not a sharp right-left partitioning of the sensory information that the respective hemispheres receive, there is much sensory information that is processed in only one of the two hemispheres. What this would seem to suggest is that, at any given time, the experiential streams of the respective hemispheres will have significant differences in content.

There is some room for the defender of the TSTS account to manoeuvre. Our eyes are constantly moving, sometimes making several saccadic movements in a second. This means that, even though the visual *experience* generated by the two hemispheres will never be the same at a time, both hemispheres may, for the most part, have very similar representations of objects and their positions in space around the individual.

¹⁴See Trevarthen and Sperry (1974) and Trevarthen (1974). Some of this research is discussed in section 4.2.1.

Whatever a given hemisphere lacks in terms of its visual experience at a time it can make up by gathering additional information when the eyes move to survey the scene. The tests that are run on split-brain individuals inside the laboratory are designed to exploit these perceptual deficits. That is why, for instance, visual cues are displayed for very short periods of time. Doing so ensures that there will not be enough time for saccadic eye movements on the part of the test subject. But outside of the laboratory, the individual is free to survey the world as he or she wishes.

The question remains, however, whether this way of making up for deficits of sensory information is enough to account for the behavioural unity that split-brain individuals demonstrate outside of the experimental conditions. It is clear that split-brain individuals may be able to compensate somewhat for the respective perceptual deficits in the two hemispheres at any given time, and this may go some way to accounting for unified behaviour at a given time or over short periods of time, but if, as the TSTS account maintains, there really are two distinct subjects of experience housed in a split brain, would it not be possible for these two subjects to slowly develop slightly diverging preferences and subtly different patterns of reasoning that would increasingly result in conflicting intentions? If so, would we not expect to see more conflicting behaviour in commissurotomy patients over time?

This objection assumes that TSTS accounts associate the subject of consciousness with the functional role of governing thought and action. There is room in logical space, however, for these to be divorced from each other. For example, one could maintain that it is possible for the content of a subject's experience to be independent, to a greater or lesser degree, from the content available to subconscious processes that are responsible for motor and linguistic outputs. The defender of a TSTS account could then attempt to deflect the above objection by locating the explanation for the behavioural unity of split-brain subjects with such subconscious processes.

Although this is a position one could take, it is not really one that is open to defenders of TSTS accounts. The main motivation for positing two streams of consciousness to begin with is the behavioural disunity exhibited by split-brain subjects in the experimental conditions. This already betrays an implicit assumption, on the part of the TSTS theorist, that the seat of experience is also responsible for thought and action. If the TSTS theorist sought to defend his or her view from the above objection by divorcing the subject of consciousness from the functional role of governing behaviour, one might wonder why, given all the problems and questions that beset views that posit multiple streams of consciousness, this strategy was not employed at the outset. Why not attempt to account for behavioural disunity at the level of sub-experiential processes instead of arguing for views that stretch our notions of subjects, conjoint phenomenology, and maximal experiential states in such awkward and problematic ways? Indeed, this is the strategy I favour and will argue for in the next chapter.

The two-stream accounts face many problems. Some of these problems could, perhaps, be tolerated if these two-stream accounts were the only options available. Some of the problems, however, are significant enough that, even if there were no other options, their respective views would remain very difficult to accept. Fortunately, there are other options available. In the next chapter, I will consider some of these options.

CHAPTER 4

ONE STREAM OF CONSCIOUSNESS

I find the views discussed in the last chapter extremely difficult to accept. Although their proponents should be lauded for attempting to respect the empirical findings from split-brain subjects, the problems and questions their views face are so deep and numerous that they should only be accepted if there are no other viable alternatives. In this chapter I will consider some alternatives. Both of the alternatives I will consider are views according to which split-brain subjects have only one stream of consciousness. Thus they are both views that are consistent with the Unity Thesis. I have already indicated that I find the Unity Thesis compelling. This is due, in no small part, to the fact that the Unity Thesis is consistent with naïve views about the structure of consciousness. In the realm of rational enquiry, however, being consistent with our naïve views only goes so far. Before a one-stream view can be accepted over any of the two-stream views discussed in the previous chapter, it must be shown that it is a better alternative.

One-stream views, by their very nature, avoid the conceptual pitfalls and conceivability questions that bedevil two-stream views. Thus, they enjoy at least one significant benefit over two-stream views. The main challenge for the defender of a one-stream view, however, is to provide an account of how a one-stream view might be reconciled with the split-brain data. If such a reconciliation is not possible, then being consistent with the Unity Thesis will not be enough to favour it over its two-stream rivals. Few have taken on this challenge. Bayne (2010) is one noteworthy exception. Since I am inclined to accept the Unity Thesis, I am interested in seeing

Bayne’s project succeed. In this chapter, I will examine Bayne’s attempt to reconcile the Unity Thesis with the empirical data. Although Bayne has gone a long way toward showing how the empirical data gathered from split-brain subjects does not ultimately undermine the Unity Thesis, I think his efforts fall short. I will present Bayne’s account and argue that it is problematic. Having shown this, I will present an alternative. Although the alternative I will present departs from Bayne’s switch model in certain important respects, it draws inspiration from it.

4.1 Bayne’s Switch Model

According to Bayne’s account, split-brain subjects have “a single stream of consciousness that switches between two streams of processing” (2010, 210). At one moment, a split-brain subject’s stream of consciousness will be supported by one hemisphere; at another moment, it will be supported by the other hemisphere. In essence, the subject’s stream of consciousness is a ball that is constantly being passed back and forth between them. This is not to say that during the time a hemisphere is not actively supporting consciousness no information processing is going on in it at all. Information processing can occur in a hemisphere even if it is not conscious at the time, but the contents of a subject’s stream of consciousness at a time are ultimately generated by only one hemisphere at a time. Bayne has fittingly dubbed his view the “Switch Model.”

The Switch Model can be thought of as a conjunction of two theses: a thesis about the nature of split-brain subjects’ conscious experience and a thesis about how a conscious experience of that nature is realized by the brain.

Split-Brain Unity Thesis The consciousness of split-brain subjects is phenomenally unified.

Switch Thesis Split-brain subjects have a phenomenally-unified stream of consciousness in virtue of the fact that only one of the two hemispheres in their brains is conscious at a time.¹

In maintaining the Split-Brain Unity Thesis, the defender of the Switch Model parts ways with accounts that posit two streams of consciousness, for the Split-Brain Unity Thesis implies that split-brain subjects have but one stream of consciousness and that their conscious mental states are always encompassed in one, total conscious mental state at a time.

There is much to like about the Switch Model. It seems to be in a good position to accommodate the split-brain data that seem to imply a fractured or disunified consciousness. If commissurotomies sever the channels whereby sensory information that is processed primarily in one hemisphere is communicated to the other hemisphere, then it would stand to reason that a subject for whom only one hemisphere was conscious at a time would act as though only one stimulus was shown when, in reality, it was one of a pair of stimuli. The model also seems to be in a position to accommodate the split-brain data that pull in the direction of a unified consciousness. Split-brain subjects can behave normally outside of experimental conditions because, even if the two hemispheres are not conscious simultaneously, they still process a lot of information simultaneously. Thus, when the ball of consciousness is passed back and forth between them, intentions and memories can be preserved.

¹Although the Switch Thesis concerns only split-brain subjects, Bayne (2010, 212) seems to maintain the view that the consciousness of normal subjects is also governed by the same switching mechanism. He cites the fact that certain split-brain phenomena can also be elicited in normal subjects. See, for instance, the findings of Milner and Dunne (1977), which I discuss in section 4.2.2, and Landis et al. (1979).

4.2 Evaluating the Switch Model

The fact that the Switch Model attempts to account for the split-brain data in a way that makes it possible to affirm the Split-Brain Unity Thesis, and, thereby, the Unity Thesis, counts in its favor. The question is whether this can be done successfully. I think it can, but not if the Switch Thesis is part of the package. In this section, I will consider a series of worries and empirical findings that call the Switch Thesis into question. First, I will discuss an objection to the Switch Thesis that Bayne himself considers. Then I will take a look at some of the studies that have been done on split-brain subjects. I will argue that the Switch Thesis is difficult to reconcile with the results of these studies and that the thesis should not be accepted.²

4.2.1 Radical Shifts

A relatively straightforward worry can be raised for the Switch Thesis by considering what the experiences of split-brain subjects would be like if it were true. If only one hemisphere were conscious at a time, would not split-brain subjects experience radical and sudden changes in the character and content of their streams of consciousness? Information from the two halves of the visual field is, for the most part, processed separately in the two hemispheres. If only one hemisphere were conscious at a time, then it would seem that a split-brain subject's visual experience would never encompass more than half of the visual field at a time. And if there were frequent switches between hemispheres in the way that the Switch Thesis maintains there are, would not the visual experience of a split-brain subject frequently switch between a visual experience of one half of the visual field followed by a sudden switch to a visual experience of the other half of the visual field? Not only would there be radical and sudden changes with respect to visual experience, would there not also be

²For another recent evaluation of Bayne's Switch Model, see Schechter (2012). Some of her criticisms are similar to some of the criticisms I raise below. This chapter was written prior to the publication of her paper.

sudden and radical changes with respect to all the other sensory modalities in which information is processed in stereo like it is in the visual system? One would expect that sudden and radical changes of this sort would not go unnoticed. Yet split-brain subjects do not report them.

Bayne (2010, 218) anticipates this worry. He attempts to deflate it by pointing out that changes with respect to the content of consciousness, even changes that are of a significant and substantial nature, are not always accompanied by an awareness of those changes. One example of this is a phenomenon known as change blindness.³ People are often oblivious to significant changes in their environment, to changes that occur between scenes in a movie, or even to changes across a series of still images. The latter is exploited in the popular and ubiquitous “spot the change” puzzles. It can often take a significant amount of time to spot the differences. Once spotted, one always wonders why the differences took so long to notice in the first place. Change blindness is also exploited by magicians, illusionists, and other performers who use slight-of-hand tricks.

Researchers who study change blindness have shown just how pervasive and deep-seated the phenomenon is. In one study (Simons and Levin, 1998), an experimenter stopped individual pedestrians to ask for directions. During each conversation, a pair of workmen interrupted by carrying a door between the experimenter and the subject. While out of view behind the door, one of the workmen carrying the door switched places with the experimenter and continued the conversation with the subject. Only half of the subjects noticed the switch of conversation partners. In another well-known study (Simons and Chabris, 1999), subjects were shown a 75-second video clip. In the clip, two teams, each consisting of three people, wearing white and black

³Bayne does not bring the phenomenon of change blindness into his discussion of the objection, but, as will become apparent immediately, the defender of the Switch Model can use it in his or her defense.

shirts respectively, walk around in random fashion passing a basketball to members of their own team. About two-thirds of the way through the clip, either a woman carrying an umbrella or a person dressed in a gorilla costume walks from one side of the room to the other through the group of people passing around the basketballs. The subjects viewing the clip were instructed to pick a team and keep a silent count of the number of passes the team made. Again, only about half of the subjects noticed the umbrella-carrying woman or the gorilla.

Many are inclined to maintain that the subjects viewing the video clip have a visual experience of the umbrella-carrying woman or the gorilla even though they do not report it. The thinking is that, if they have visual experiences of the people passing around the basketball, they must also have a visual experience of the umbrella-carrying woman or the gorilla. The same would be the case in the situations involving the workmen carrying the door. Surely the change in conversation partner results in a corresponding change in content of the subject's visual experiences. Although the person being asked for directions is not aware of the change, the change must be real. If normal subjects in situations like this can be subject to change blindness, could not split-brain subjects also be oblivious to the radical changes in their experience?

These cases involve normal subjects. But there are other examples, involving non-normal subjects, that make the point even more strongly. Bayne points out that patients who suffer certain kinds of brain damage and lose their ability to see color are often unaware that they have lost the ability. Subjects with unilateral or hemi-neglect manifest this lack of second-order awareness in an even more pronounced fashion; they lack visual awareness of objects in half of their visual field. They have been known to eat the food on only one half of their plate or to shave only one side of their face. In another oft-cited study (Bisiach and Luzzatti, 1978), two subjects with hemi-neglect were asked to imagine standing on one end of Piazza del Duomo in Milan, a place they knew well, and describe what they saw. The subjects only named

objects that would have been located to their right and failed to mention any that would have been on their left. When asked to imagine standing on the opposite end of the Piazza facing their previous location, they again only named objects on their right. Both subjects seemed completely oblivious to the fact that they had named non-overlapping sets of objects.

The phenomenon of change blindness shows that subjects, even normal ones, can be unaware of significant changes to the contents of their experiences and the data gathered from subjects suffering from unilateral or hemi-neglect show that it is possible for a subject to be unaware of rather significant experiential deficits.⁴ The question here is whether these things can serve as an explanation for why split-brain subjects do not report the significant changes in their experiential lives that the Switch Thesis seems to call for.

One question that seems relevant to deciding the adequacy of this as an explanation for why split-brain subjects do not report the switches between their sensory half-fields is the question regarding the frequency with which these switches would occur. Subjects suffering from hemi-neglect live with a perpetually compromised sensory field. One might think of their experience as being like that of a split-brain subject whose experiential stream never switches from one half-field to the other. Since it is possible for such a subject to be completely unaware of his or her compromised sensory field, it might not be such a stretch to think that, if a split-brain subject's stream of consciousness switches between hemispheres relatively infrequently, the switches could go unnoticed. If the switches occur frequently, however, this explanation for why split-brain subjects do not report them seems much more difficult to believe.

⁴Many subjects suffering from unilateral or hemi-neglect have largely intact and properly functioning sensory systems. Thus, it is an open question whether, strictly speaking, they fail to have sensory experiences of objects in the affected half-field or merely fail to incorporate the sensory information from the affected half-field into their voluntary actions and verbal responses.

Is there any reason to think that switches between hemispheres could not occur relatively infrequently rather than frequently? One might have the following worry. If the switches only occur relatively infrequently, would not split-brain subjects have much more difficulty navigating their environment? They would go for longer periods of time without being aware of objects within one of their sensory half-fields. Would they not stumble into objects in their environment more frequently as a result? Perhaps not. Split-brain subjects are still able to move their eyes and turn their head like normal subjects, and so they would still be able to gain much of the information about their environment they would need to be able to navigate it effectively.

Does this explanation for why split-brain subjects do not report dramatic and sudden changes in their stream of consciousness discharge the worry for the Switch Thesis? Not entirely. The reason for this is that the data permit more than one interpretation. Split-brain subjects do not report the kind of large scale experiential changes one would expect them to undergo if the Switch Model were correct, but the absence of these reports is compatible with two distinct possibilities: one according to which split-brain subjects undergo the changes to their experiences the Switch Thesis implies they should undergo but they are unaware of those changes, and another according to which the experiences of split-brain subjects do not actually undergo the kinds of changes the Switch Thesis predicts. The above considerations involving change blindness and hemi-neglect are enough to make the first of these two possibilities a live one, but they do not rule out the second. Indeed, the alternative to the Switch Model I will outline in section 4.3 takes the latter possibility to be actual.

There is another thing the defender of the Switch Thesis can say by way of response to the worry about radical shifts, and that is to argue that the experiential shifts would not be as radical as one might be inclined to believe. The simplistic view of how sensory information is parcelled to the two hemispheres is that it is split evenly down the middle; visual information from the left half of the visual field goes to the

right hemisphere, visual information from the right half of the visual field goes to the left, and sensory information from the other senses is split down the middle in the same way. Studies on the peripheral vision of split-brain subjects has shown, however, that, at least in the case of vision, the hemispheres receive some visual information from both visual fields.

Trevarthen and Sperry (1974), building on some earlier work by Trevarthen (1970), designed a set of tests for four split-brain subjects.⁵ The subjects were asked to fixate on a point in the centre of their visual field. Then they were shown various stimuli on the outer halves of their visual field and asked to report on them in various ways. In one set of trials, a white circle was shown in each visual field. Some of the subjects were able to report when both of the circles moved up, when both of them moved down, and when one moved up and the other down. In another set of trials, some of the subjects noticed when the circles expanded or contracted. In another test, the subjects were told to indicate when a line that was slowly rotating about the fixation point in one visual field lined up with a line that was stationary in the other visual field. Some of the subjects were quite successful in this task. In yet another set of trials, some of the subjects could describe the speed and direction of motion of dots in their peripheral vision. Two of them could count up to three or four squares correctly. All of the subjects could correctly report a change in colour of the objects in their peripheral vision, though there was some disparity between them in their ability to name the colours they were shown.

These studies show that split-brain subjects are capable of some, admittedly rudimentary, visual discriminations in the peripheries of their visual fields. Split-brain subjects may not be able to verbally identify objects that are shown to them nearer the midline in the left half of their visual field, but they are capable of responding

⁵For a more recent survey of visual integration in split-brain subjects, see Corballis (1995).

verbally in visual discrimination tasks that involve the periphery of their left visual field. More importantly, they are capable of making certain comparisons between stimuli that are shown in the peripheries of both halves of their visual field. Thus, it would seem as though split-brain subjects have visual experiences that encompass both halves of their visual field simultaneously.

The surgical procedures that disconnect the two hemispheres of the brain from each other leave the mid- and lower-brain structures intact. The thinking is that some visual information travels between the hemispheres through these sub-cortical channels. Neuroscientists distinguish between the primary and secondary visual system. In general terms, the primary visual system is what permits us to see objects in fine detail and to recognize them. The secondary visual system, on the other hand, seems to be more involved in spatial integration and plays an important role in directing eye saccades and facilitating large-scale movement. In primitive vertebrates, the equivalent of the secondary visual system is, in fact, the principle visual system. These two visual systems take different paths through the brain and have been shown to be somewhat independent from each other. Since both of these visual systems go through mid-brain structures that are left intact by the commissurotomy, it is quite possible that some visual information crosses the vertical mid-line in the process. If this is the case, the hemispheric switches the Switch Thesis predicts would not result in experiential changes that are as radical as one might initially believe.

These considerations go some distance to addressing the question why split-brain subjects do not report the kind of radical shifts one might expect them to experience according to the Switch Thesis. There are other considerations, however, that are much more difficult for the defender of the Switch Thesis to account for. These will be my focus in the next sections.

4.2.2 Chimera Studies

A lot of what we know today about the differing specializations of the two hemispheres of our brains we know, in part, as a result of research that was conducted on split-brain subjects in the 60s and early 70s. Knowledge of these specializations was gained by recording the performance of one hemisphere on a given task, and then doing the same for the other hemisphere. Different tasks revealed the different specializations. Chimera studies were designed to test what would happen when the two hemispheres and their respective specializations were played off of each other.

In a typical chimera study, split-brain subjects are shown images that are constructed by taking halves from two separate pictures or line drawings and combining them into one at the vertical midline. In some studies, these chimeric images are constructed by taking photographs of the faces of different people, cutting the photographs at the vertical midline, and recombining them so that each image contains half of one person's face on the one side and half of another's face on the other. In other studies, the chimeric images are recombinations of line drawings of various ordinary objects. The subjects are told to look at the fixation point in the experimental apparatus and then they are shown a chimeric image for about 150 milliseconds—this short exposure ensures against eye saccades. Their instructions are to pick, from an array of drawings or pictures on the table before them, one that matches the image they see.

In one oft-cited study of this sort, Levy and Trevarthen (1976) explored how different instructions to the subjects would effect the response they gave. Sometimes the individuals were instructed to select the object that *looked* similar to what they saw, sometimes they were told to select the object that one would *use with* the object they saw, and sometimes the individuals were given an ambiguous instruction like “select the one which is similar to, goes with, or is like in some way the one you see” (ibid., 302). That is, sometimes the subjects were told to match via appearance, a

right-hemisphere specialization, sometimes via function, a left-hemisphere specialization, and sometimes, as when they were given the ambiguous instruction, they were not instructed what the match was to be selected for.

What they found was that, in the trials with unambiguous instructions, the subjects selected the object that was shown to the hemisphere that was specialized for the task. For example, if the subject was shown an image with a pair of scissors on the left and a set of eye glasses on the right and instructed to match for appearance, the subject would select an object that matched the scissors rather than the eye glasses in appearance. If the subject was shown the same image and instructed to match for function, the subject would select an object that matched the eye glasses rather than the scissors in terms of function.

There are two things of significance to take note of with respect to the Switch Thesis. First, even though a different object was shown to each hemisphere simultaneously, the subjects responded by selecting only one object as a match. In a later paper in which she generalizes about her observations on numerous different studies like this, Levy (1990, 235) reports that, not only did the subjects select only one object as a match, they gave no indication that they were aware that they had been shown a chimeric image. The non-responding hemisphere did nothing to show that it disagreed with the selection that was made. If the selection was an appearance-match made by the right hemisphere, the left hemisphere did not verbally renounce or distance itself from the selection that was made. Likewise, if the selection was a function-match made by the left hemisphere, the right hemisphere did nothing to indicate disagreement with the selection.

Similar results have even been obtained in normal subjects. Milner and Dunne (1977) found that, if the two halves of a chimeric image are separated by a band of white, normal subjects can participate in numerous trials before becoming aware that they are being shown chimeric images. Nearly half of the subjects became aware of

the fact they were being shown chimeric images, but only after having seen a dozen or more of them. The rest of the subjects in their study never became aware that they were being shown chimeric images at all.

The second thing to take note of is the conclusion that Levy and others have drawn from these results. The results seem to indicate that there is some sort of metacontrol mechanism in the brain that determines which hemisphere will respond in a given trial. When the instruction is to match for appearance, the right hemisphere registers an appropriate response, and when the instruction is to match for function, the left hemisphere registers an appropriate response. Thus, it can happen that the very same chimeric image can generate a different response depending on which instruction is given before the image is shown.

These results seem to confirm the Switch Thesis. Given that only one hemisphere responds and that the other shows no disagreement, these results would suggest that only one of the two hemispheres is conscious at a time. And given the effects that the instructions prior to the chimera trials have on the kind of response that is given, it would seem that these studies provide evidence of a mechanism in the brain that oversees the switching of consciousness between the hemispheres that the Switch Thesis posits. Later, when I present an alternative to the Switch Model, I will offer a rival interpretation of these results. What I will do here, however, is show that, when the results of the chimera studies are taken in their entirety and examined more closely, they do not confirm the switch thesis as well as its defenders think it does. In fact, some of the results from chimera studies seem to undermine the Switch Thesis. This is significant, because Bayne, who, to my knowledge, has developed the most thorough defense of the Switch Model to date, employs chimera studies in motivating his view.

Given the fact that the non-responding hemisphere offers no rebuttal or registers no disagreement with the responding hemisphere's selection, the natural conclusion

to draw is that the non-responding hemisphere is non-responding in virtue of the fact that it is unconscious or, as Levy says, that “no perception [is] constructed by [it]” (1990, 235). In an attempt to test this conclusion, she and her colleagues (Levy et al., 1972) constructed a set of chimera trials in which subjects were occasionally interrupted and asked to respond to the stimulus in a different way than they were initially instructed to do. Subjects were instructed to point to the matching image from an array of images in front of them. Occasionally, the experimenter would, just before the subject made a manual match, remove the images and tell the subject to verbally describe the image he or she had just seen instead. Most of the time, the subject would then describe the face that had been shown to the left hemisphere. When the experimenter then reproduced the array of images and asked the subject to make a manual match, the subject usually proceeded to make a match based on the face that was shown to the right hemisphere. Interrupted trials like this were also run in the inverse order, instructing the subject, first, to describe the face verbally, and then interrupting and instructing the subject to make a manual match. Occasionally, the subjects realized the inconsistency between their two responses and evidenced confusion.

Similar observations have been made in experimental situations that do not make use of chimeric images but are analogous to chimera studies in other ways. For instance, Sperry (1968a) simultaneously flashed different symbols to the opposite hemispheres of a split-brain subject; the right hemisphere was shown the dollar sign and the left was shown a question mark. With his hands shielded from sight behind a barrier, the subject was asked to draw, with his left hand, what he saw. He drew a dollar sign. When he was then asked to verbally report on what he had just drawn, he indicated that he had drawn a question mark, the symbol that had been shown to his left hemisphere.

These observations call the Switch Thesis into question because they seem to suggest that, when visual stimuli are flashed to both hemispheres simultaneously, each hemisphere is visually conscious of the stimulus that is flashed to it. How else could each hemisphere accurately match or describe the stimulus that is shown to it? And if both hemispheres are conscious of their respective visual stimuli when they are flashed, then the Switch Thesis, which maintains that only one hemisphere is conscious at a time, would seem to be false.

This conclusion also seems to be confirmed by some studies that have been conducted on split-brain monkeys. Trevarthen (1962), for instance, showed that the two hemispheres can learn simultaneously. Monkeys were placed in front of a pair of translucent screens. A pair of projectors placed behind the pair of screens then flashed contradictory images onto the screens. One projector flashed a circle onto the left and a cross onto the right. The other projector did exactly the opposite. Polarizing filters were placed between the projectors and the screens and also between the screens and the monkey's eyes so that, when one hemisphere saw cross-circle (a cross on the left and a circle on the right), the other saw circle-cross. The stimuli were simultaneously flashed onto the two screens and the monkey was rewarded for pushing the appropriate screen in response. In this way, over a series of trials, one hemisphere learned that pushing crosses was rewarded while the other hemisphere learned that pushing circles was rewarded. How could the two hemispheres simultaneously learn to respond to different stimuli unless they were simultaneously conscious when the stimuli were being shown?

It is open to the defender of the Switch Thesis to respond by maintaining that the two hemispheres of split-brain subjects are alternately conscious during these interrupted trials, depending upon which hemisphere is being asked to respond. The defender of the Switch Thesis could say that, when the experimenter interrupts and asks for a description of the flashed image, the left hemisphere is conscious and

responds, and then, when the array of images is put back in front of the subject and the subject is asked to manually select a match, the right hemisphere becomes conscious and responds. In this way, it could be that only one hemisphere is conscious at a time even though it would be natural to conclude from the observations that both were conscious simultaneously.

This way of accounting for the data is, I think, problematic. Because subjects of interrupted chimera trials do not give their responses simultaneously, the Switch Thesis seems to be able to account for the data. However, what cannot be overlooked is that the stimuli are *given* to the two hemispheres simultaneously. If, as the Switch Thesis maintains, only one hemisphere is conscious at a time, then one of the hemispheres will be unconscious at whatever time the stimulus is flashed. This means that, even though both hemispheres are later able to manually match or describe, depending on the hemisphere, the image that was flashed, one of them will be doing so in spite of the fact that it was unconscious when the image was shown.

Bayne allows that processing can continue to take place in the unconscious hemisphere. Could the Switch Thesis be saved by maintaining that unconscious processing in the non-conscious hemisphere is sufficient to generate the responses that are given in interrupted chimera trials? Nobody will deny that our behavior can be influenced by signals that our sensory organs respond to but which we are not conscious of. Nevertheless, the way in which we react to and are influenced by unconscious sensory information is very different from the way in which split-brain subjects are able to respond to visual information that, according to the Switch Thesis, is received unconsciously. The right hemisphere is capable of pointing out an image that matches in appearance the object it was shown. The left hemisphere is capable of describing the object that it was shown. Subjects do not generally act on unconscious sensory information in these ways. In fact, it is very nearly *definitional* of the notion of consciousness that unconscious mental states do not manifest themselves in such be-

havior. The act of pointing at an object could, in certain circumstances, be a response to unconscious sensory information. But what differentiates the act of pointing in these interrupted chimera trials is that the pointing is discriminatory. It is not a mere reflex, nor is it random or guesswork. The verbal descriptions of the flashed images are, needless to say, even more indicative of consciousness.

It is, perhaps, worth pressing this line of attack a bit further. In normal, uninterrupted chimera trials, experimenters report that the subjects respond to the images with “placid equanimity” (Levy, 1990, 237); they respond as though nothing out of the ordinary has happened. However, during interrupted trials, something significantly different happens. After the second response has been given, a response that inevitably conflicts with the first, the subjects often manifest “considerable perplexity and confusion” (Levy et al., 1972, 67). The question for the Switch Thesis is this: if one hemisphere is unconscious while the other gives its response, why would the subjects manifest any perplexity or confusion? It would seem that confusion or perplexity would only be warranted if one or other of the two hemispheres was aware that its response conflicted with the response given by the other hemisphere. And how would such an awareness arise if each hemisphere was unconscious while the other responded?

The defender of the Switch Thesis seems to be in the same position he or she was in earlier. The fact that the two hemispheres are capable of responding successively to images that are shown to them simultaneously would seem to indicate that split-brain subjects are capable of responding to unconscious visual information in ways that are normally indicative of consciousness. The fact that these subjects then manifest confusion after they have given the second, conflicting response would suggest more of the same. The defender of the Switch Thesis must say that, when one hemisphere responds, the other receives the information regarding the response

but is not conscious of it. Then, when it, in turn, becomes conscious, it can manifest confusion even if it was not conscious of the confusion-inducing stimulus.

Defenders of the Switch Thesis have a way of accounting for the way in which split-brain subjects respond to interrupted chimera trials, but it requires them to maintain that these subjects are capable of responding to unconscious visual information in ways that are normally thought of as sure indicators of consciousness. Perhaps this consequence of the thesis is one that its defenders are prepared to swallow. I think it is not worth the trouble.

4.2.3 Simultaneous Dual Voluntary Responses

There are experimental results that are even more difficult to reconcile with the Switch Thesis. So far, I have been pressing the Switch Thesis by pointing to empirical findings that would seem to indicate that the two hemispheres are capable of simultaneous conscious awareness. The empirical findings I will draw attention to now seem to indicate that the two hemispheres are capable of simultaneous voluntary action.

There has been considerable discussion about the degree to which the attentional mechanisms of split-brain subjects are linked across the two hemispheres. Normal subjects are able to respond more quickly to stimuli that appear where their attention is focused and more slowly to stimuli that appear elsewhere. This also holds true for split-brain subjects, but, interestingly enough, this dynamic does not particularly respect visual hemifield boundaries (Corballis, 1995). When split-brain subjects are presented with a cue in a certain region in one half of their visual field, they respond more quickly to a stimulus that is then presented at that location than at another elsewhere in that visual field, just like one would expect. However, they seem to respond just as quickly if the stimulus is presented at the cue's corresponding location in the opposite visual field; the fact that the stimulus appears in the opposite visual field does not seem to detract from the subject's reaction time. This implies that

the attentional mechanisms of the two hemispheres are linked even for split-brain subjects.

Bayne (2010, 214–216) takes this data to be an objection to the Switch Thesis. He responds by invoking experimental results that seem to show that the attentional mechanisms of the two hemispheres in split-brain subjects are somewhat independent. For instance, he points out that split-brain subjects seem to perform better than normal subjects in tasks that require them to visually track more than one object at a time. It is not clear to me why Bayne would take these results to support the Switch Thesis. The reason for this is that visual tracking seems to presuppose consciousness. If there is reason to suppose that each hemisphere is capable of visually tracking an object, and if they are capable of visually tracking their respective objects simultaneously, then it would seem that both hemispheres are simultaneously conscious.

One might question this conclusion by pointing to situations in which visual tracking seems to be occurring unconsciously. One might recall Armstrong’s (1981) long-distance truck driver. Many of us, regrettably, have first-hand experience with the phenomenon of driving in an absent-minded state. After driving for a while, perhaps late at night, on a road that one knows well, one can “come to” and realize that one has no memory of the last mile. Yet even though one’s mind has been focusing on other things or, perhaps, has been in a quasi-wakeful state, it is clear that some attentional processes have been monitoring the road and the traffic. Otherwise, there would have been a collision.

I do not think that this familiar phenomenon is a case in which visual tracking is occurring independently of consciousness. Of course, defenders of the Switch Thesis who might offer this as evidence of the possibility of unconscious visual tracking will not insist that drivers who are lost in thought are unconscious in different sense, i.e., in the way that a sleeping person or person under anaesthesia is unconscious. All

they need to claim is that the mental processes that are keeping the absent-minded driver on the road are unconscious. This, if true, would open the possibility that only one of the hemispheres is conscious at a time in independent tracking trials. But the reason the absent-minded driver is not a case in which visual tracking is occurring independently of consciousness is that, even though the driver may be thinking about something else and not focusing on what is happening on the road, the driver is still the subject of a visual experience. It is not as though one's visual experience is extinguished when one becomes lost in thought or when one's attentional resources are focused elsewhere. Thus, the fact that one is capable of visually tracking objects without, at the same time, focusing on them does not show that visual tracking can happen independently of consciousness.

But even if the facts about absent-minded drivers are different than I believe them to be, there is a way in which what occurs in the independent-tracking studies is relevantly disanalogous to what occurs with absent-minded drivers. Split-brain subjects, and normal subjects for that matter, are visually attending to the stimuli they track in independent-tracking studies. Visually tracking multiple objects requires focus. It requires a lot of focus. Absent-minded drivers, on the other hand, are not focused on what their eyes are tracking. The subjects of independent-tracking tasks are summoning all of their attentional resources and, more importantly for my purposes here, are voluntarily controlling what their attentional resources are directed at. This is significant, because, if the reason split-brain subjects can perform better on some visual tracking tasks than normal subjects can is that their two hemispheres are capable of working more independently in some situations than normal brains are, then it would seem that both hemispheres are voluntarily controlling what they are attending to and are doing so simultaneously. If voluntary control requires consciousness, then these visual-tracking studies undermine the Switch Thesis.

These facts about the way in which the attentional mechanisms of split-brain subject work manifest themselves in numerous ways. Consider a study done by Gazzaniga and Sperry (1966), in which they compared normal and split-brain subjects with respect to their performance on a double-discrimination task. Subjects sat facing a screen, fixating on a point at its center. Their right and left index fingers were then placed in a ready position, touching the screen a small distance from the fixation point on either side respectively. The discrimination task on the left side was between white and black, whereas the discrimination task on the right side was between red and green. The subjects were instructed to respond to stimuli as quickly as possible by moving their respective index fingers to the right or left of their respective ready positions. What Gazzaniga and Sperry found was that it took significantly longer for normal subjects to respond on trials in which the two discrimination tasks were given simultaneously than in trials in which only one discrimination task was given. Split-brain subjects, on the other hand, did not take significantly longer to respond when the two discrimination tasks were given simultaneously. A natural way to understand the results of this study would be to attribute the lack of a significant increase of time for the split-brain subjects in the simultaneous double-discrimination tasks to the fact that the two tasks are completed simultaneously. How else could it be that it takes no longer for a split-brain subject to complete two discrimination tasks than it does to complete one?

Another study, this one conducted by Franz et al. (1996), shows a similar sort of disconnect between the performance of normal subjects and a split-brain subject on bimanual tasks. The subjects in this study sat with their hands shielded from view by a paper screen. A writing device was placed in each hand. Simple figures—in one set of trials, straight lines, and in another set of trials, three-sided rectangles—were then flashed into one or both hemispheres at the same time. The subject was instructed to then reproduce what had been flashed. When a stimulus was flashed to only one

hemisphere, all of the subjects completed the task with ease. Similarly, when a pair of stimuli of the same orientation and size were flashed to the two hemispheres simultaneously, all of the subjects were able to complete the drawings with relative ease. Interesting and significant differences emerged, however, when stimuli of differing sizes and orientations were flashed to the two hemispheres. On these trials, the split-brain subject performed significantly better than his normal counterparts.

Like the results in the study by Gazzaniga and Sperry, the results in this study also undermine the Switch Thesis. In this study, even when performance on the tasks is relatively poor, both hands are voluntarily moved simultaneously. What could be a better indicator of consciousness than voluntarily controlling a limb in reproducing a visually presented figure? Perhaps the defender of the Switch Thesis could maintain that the reason normal subjects are able to draw two figures of similar size and orientation with ease is that elements of such a task can be completed with a relatively high degree of automation. Then, if a switch mechanism were in place, it could switch consciousness from one hemisphere to the other without a significant effect on the performance of the unconscious hemisphere. The poorer performance on trials in which the stimuli are of different sizes and orientations could be explained by the fact that the task, in such trials, is too difficult to complete without a significant degree of conscious oversight. Although the results of the trials for normal subjects might be accounted for in this way, it is difficult to see how the Switch Thesis could account for the fact that the split-brain subject was able to complete trials involving stimuli of differing sizes and orientations without a significant decrease in performance. If such a task requires a significant degree of conscious oversight, then split-brain subjects should also perform poorly on it.

There is more evidence that the two hemispheres are capable of engaging in voluntary behavior simultaneously. The testing strategy in some of Sperry's classic tests consisted in exposing one hemisphere to an object, either by revealing it in one half of

the visual field or by placing it in one of the subject's hands, and then requiring the non-exposed hemisphere to identify it verbally by naming it or manually by selecting it from a collection of objects. In a departure from this sort of task, Sperry placed a distinct object in each hand of a split-brain subject and then removed and hid the objects in a scrambled pile of test objects. Sperry (1968a, 727) reported that the subject hunted through the pile simultaneously with both hands until each hand found the object it had previously held. Sometimes one of the hands would pick up and examine the object the other had been given only to discard it and go on searching for the object it was looking for. The simultaneous searching behavior of the two hands in this experimental situation would again seem to indicate that the two hemispheres are capable of simultaneously exerting voluntary control over the two limbs and, thus, that they are simultaneously conscious.

It is often conceded that there is more overlap of motor control over the limbs than a simple left-right division would allow for. This said, an interpretation of simultaneous manual searching according to which both hands are under control of the same hemisphere is not really an option here. Much has been made over the fact that when an object is given to the one hand to hold, the ipsilateral hemisphere remains in the dark about it. If it was maintained that a single hemisphere was exerting motor control over both hands in the simultaneous searching tasks, it would be completely inexplicable how the ipsilateral hand would eventually find the object it had previously been given to hold. As long as we are working under the received framework in which the motor control of each hand resides in the contralateral hemisphere, it would seem that simultaneous search tasks undermine the Switch Thesis.

There are even more remarkable cases in which the two hemispheres seem to be exerting simultaneous motor control over their respective limbs. In a situation described by Mark (1996, 192), a split-brain subject was asked if she was still having

seizures.⁶ When she tried to enumerate them, each hand held up a different number. In another scene described by Wilkes (1978, 187), a split-brain subject was asked to build a structure out of building blocks with the right hand. The left hand kept on seeking to interfere until the subject eventually sat on it, leaving the right hand to proceed uninterrupted. In another study, one intended to test the self- and social awareness of the two hemispheres, Sperry et al. (1979) outfitted their subjects with contact lenses that occluded the right half of their visual field. This piece of equipment made it possible for the subjects to look at images for as long as they wished without visual information being fed to the left hemisphere. One of the subjects, LB, was shown an array of four pictures. Three of the pictures contained people that were unknown to him. The fourth was a picture of Hitler, in uniform, standing with several other men. LB was allowed to examine the array and asked if he recognized anyone. After about a quarter of a minute, LB pointed to the picture of Hitler. The following exchange then ensued between the examiner and LB.

Ex: “Well, on this: is this one a ‘thumbs-up’ or a ‘thumbs-down’ item for you?”

LB: Signaled “thumbs-down”.

Ex: “That’s another ‘thumbs-down’?”

LB: “Guess I’m antisocial.” (Because this was his third consecutive “thumbs-down” [response].)

Ex: “Who is it?”

LB: “GI came to mind. I mean . . .” Subject at this point was seen to be tracing letters with the first finger of the left hand on the back of his right hand.

Ex: “You’re writing with your left hand; let’s keep the cues out.”

⁶For a more extended discussion of these cases, see section 2.3.1.

LB: “Sorry about that.” (ibid., pp. 159–160)

With enough leading questions, LB was eventually able to verbally identify the person in the photograph as Hitler. Sperry and his colleagues report that it was a “standard trick” with LB to attempt to cue the verbal hemisphere in this way. Although this is not an example of simultaneous voluntary goal-directed hand movement, it certainly is an example of a case in which both hemispheres seem to be simultaneously attempting to complete a task. The left hemisphere is attempting to name the object that was indicated by the right hemisphere and the right hemisphere is trying to cue the left as to the identity of the object.

Although there are ways in which the defender of the Switch Thesis can respond to some of the objections and empirical research I have discussed, I think that the thesis ultimately falls. There is considerable reason to believe that both hemispheres can be, and often are, simultaneously conscious.

4.3 An Alternative to the Switch Model

The following question presents itself: if the Switch Thesis and, correspondingly, the Switch Model are false, then how can the Unity Thesis be maintained? In this section, I will present an alternative to the Switch Model that is consistent with the Unity Thesis. It has something in common with the Switch Model in that it affirms the Split-Brain Unity Thesis. However, it departs from the Switch Model in that it rejects the claim that only one hemisphere can be conscious at a time.

The Switch Model was motivated, in large part, by the data collected from split-brain subjects in chimera studies. Chimera trials, you will recall, seem to imply the existence of a metacontrol mechanism that dictates which hemisphere will respond to a given stimulus. This is because the subjects in these trials offer only one response to the chimeric images even though the images present two possible appropriate responses. Moreover, which of the two possible responses a subject will offer on any

given trial seems to depend somewhat on the instruction he or she is given beforehand; the instruction seems to prime the subject in a certain way that favours a response to one half of the chimera over the other. Because of this priming effect, the same subject can respond in different ways to the same chimeric image when that image is shown in sets of trials in which different instructions have been given to the subject beforehand.

The Switch Model incorporated the notion of a metacontrol mechanism into its theory of consciousness. According to the Switch Model, the alternating way in which split-brain subjects respond to chimeric images is a good guide to understanding the structure of their conscious experience; whichever hemisphere is responsible for the given response of a split-brain subject is the hemisphere that is conscious at that time. Since the responses seem to come from only one hemisphere at a time, the Switch Model seemed to fit the data nicely. As we saw, however, it does not fit *all* of the data. Nevertheless, I think there is much about the Switch Model and its posit of a metacontrol mechanism that can be of service to a theory of consciousness.

The metacontrol mechanism is, first and foremost, a mechanism that governs action, i.e., the responses of the subjects. But theorists have gone further and suggested that the mechanism may have a crucial role to play in perception and consciousness. Levy (1990, 231), for instance, argues that the results of the chimera studies confirm Sperry's view that perception is the preparation for response. On this kind of view of perception, the metacontrol mechanism implied by the chimera studies is responsible not only for the actions and responses that the subjects of the studies give but also for perception. From this, it is but a short further step to maintain that the metacontrol mechanism also governs consciousness itself. Indeed, Bayne employs the work of Levy in motivating the Switch Model. I think that there is room to divorce the metacontrol mechanism suggested by the chimera studies from the role of governing consciousness, and this is how I will attempt to reconcile the alternative view I am

developing here with the empirical findings that seem to imply that consciousness is fractured or disunified in some way. If the role of the metacontrol mechanism with respect to generating conscious experience is significantly limited, then, I will suggest, we are left with an account that is both plausible and consistent with the empirical data.

The picture is something like this. The two hemispheres together support a single stream of consciousness. This happens because the sensory systems of the two hemispheres work together to populate a single phenomenal field. The metacontrol mechanism is ultimately responsible for all voluntary action. It monitors the sensory systems and dictates which motor systems will be engaged to respond. The systems under the oversight of the metacontrol mechanism are also consumers of sensory information. In normal subjects, these systems may have access to a relatively large percentage of the sensory information that is made phenomenally manifest. But in split-brain subjects who have had their corpus callosums severed, these systems have access to a much smaller percentage of the sensory information that is made phenomenally manifest. Furthermore, the metacontrol mechanism, although it itself has access to sensory information, does not impart sensory information to the systems it governs. It merely prompts them to respond to the information they themselves may or may not have access to. Thus, there is not a one-to-one correspondence between the sensory information that is made phenomenally manifest for a subject and the sensory information that is consumed by any one of the various systems under the purview of that subject's metacontrol mechanism. So for instance, a stimulus in such a subject's left visual field might be phenomenally manifest to that subject without information about that stimulus being available to the system responsible for speech production. Such a subject would then have an experience of the stimulus without being able to verbally identify it. The system responsible for controlling the movement of the left hand, however, might well have access to information about the stimulus

and, thus, be able to manually identify it when prompted by the metacontrol system to do so. In this way, many of the classic observations of split-brain subjects can be accounted for. The numerous cases in which a sensory stimulus is presented to one hemisphere but not another, resulting in certain inabilities on the part of the subject to make appropriate matches or identifications, can be explained in this way. Even though the subject may have an experience of the stimulus, the reason the subject cannot respond in the appropriate way is because the system that has been prompted to respond does not have access to the information it needs in order to respond correctly. To use a distinction that was introduced by Block (1995), information from a given sensory stimulus may be phenomenally conscious for the subject but may fail to be access-conscious for all of a subject's response systems.

This account is also able to account for the results of the interrupted chimera trials. The interrupted chimera trials, you will recall, presented difficulties for the Switch Model because, if only one hemisphere can be conscious at a time, it would seem that the responses of subjects in interrupted chimera trials would have to be responses to stimuli the subject was not conscious of. According to the picture I have just described, the behavior of split-brain subjects in interrupted chimera trials is much more easily accounted for. According to it, both hemispheres are conscious at the time the chimeric image is shown. Thus, when the subject is interrupted and instructed to match the image according to a different criterion, the subject is able to provide a match, because the responding hemisphere was conscious at the time the image was presented. The metacontrol mechanism is not charged with bringing the responding hemisphere online. All the mechanism does is govern the response of the subject.

One more thing needs to be said about the metacontrol mechanism. In arguing against the Switch Model, I called attention to many cases in which the hemispheres seem to be giving simultaneous responses. Can this account accommodate these

cases? It can, but it needs to be stipulated that the metacontrol mechanism is capable of engaging systems in both hemispheres simultaneously. Perhaps this stipulation is entirely uncontroversial. However, the metacontrol mechanism is usually discussed in relation to contexts, such as the chimera trials, in which only one hemisphere responds at a time. This makes it easy to think about the metacontrol mechanism as one that must always choose *between* hemispheres. It is clear, however, that there are times in which motor (and other) systems in both hemispheres are active simultaneously. If all of these systems depend on the metacontrol mechanism to be activated, then the mechanism must be capable of engaging systems in both hemispheres simultaneously.⁷

The view I have just outlined does a better job of accounting for the empirical data than the Switch Model. Thus, it enjoys a theoretical advantage over them. Since the view I have outlined also accounts for the structure and nature of split-brain subjects' conscious experience in terms of a single, phenomenally unified stream of consciousness, it avoids the problems faced by the two-stream and partial-unity accounts. Thus, it also enjoys a theoretical advantage over them. Before concluding I will consider some potential worries for the view. None of these worries, in my view, present serious problems for the view, and so I think they do not ultimately shift the cost-benefit analysis in favor of any of its competitors.

Objection 1 This account maintains that there is not a one-to-one correspondence between the sensory information that is made phenomenally manifest for a (split-brain) subject and the sensory information that is consumed by the various systems

⁷It is also possible that there are multiple meta-control mechanisms at work simultaneously. Although I have characterized my alternative to the Switch Model as positing just one meta-control mechanism, it would not be opposed to the spirit of my alternative to posit additional meta-control mechanisms. The essential feature of my account is that the content of experience need not correspond perfectly to the contents that the systems responsible for behaviour and under the control of the meta-control mechanism (or multiple meta-control mechanisms) have access to. I have merely found it convenient to formulate my alternative in terms of a single meta-control mechanism because of its dialectical position with respect to Bayne's Switch Model.

involved in a subject's response to stimuli. Does this not suggest that there is a system in the brain whose sole function is to generate phenomenal states? And if so, is this not a problem? If the system that generated phenomenal states did nothing else but generate phenomenal states, it would seem to present no survival advantage and, thus, would not be amenable to evolutionary explanation.

The first thing to point out is that, as long as such a system did not come with a significant survival *cost*, it would not, strictly speaking, be incompatible with evolutionary explanations. The second, and perhaps more important, thing to say is that the account does not require that the system responsible for generating phenomenal states be functionally limited and isolated from other systems in the brain in this way. It could turn out that the system that is responsible for generating phenomenal states is the same system that makes sensory information available to other systems such as those that are responsible for speech production, movement of the limbs, etc. It could even be that sensory information is made conscious in virtue of being consumed by motor systems and other systems involved in intentional action. All the model I have presented is committed to is the claim that it is possible for a system under the purview of the metacontrol mechanism to be limited in its access to sensory information, that no single consumer of sensory information is a consumer of all of the sensory information that is phenomenally manifest for a subject at a time. But this is still consistent with the claim that there is a one-to-one correspondence between the bits of sensory information that are phenomenally manifest for a subject at a time and the bits of sensory information that are consumed by *some consuming system or other*.

Objection 2 According to this account, the norm is for both hemispheres to be conscious simultaneously. If this is so, then why, in so many situations in which separate stimuli are presented to the two hemispheres, do split-brain subjects only give a single response and give no indication that they perceived the second of the

two stimuli? Should not the split-brain subjects in these situations give conflicting responses much more often?

If the theorists who posit the existence of the metacontrol mechanism are correct, then many, if not all, of the voluntary responses of split-brain subjects are governed by this mechanism. The facts about how this mechanism works are contingent facts that can only be known *a posteriori*, by conducting inquiries that will reveal how it operates. This account could well get the contingent facts about this mechanism correct. There is nothing about the fact that a particular hemisphere is conscious that necessitates that every response system housed within it must respond to every stimulus that it is conscious of. Similarly, the mere fact that some consuming system is fed some sensory information does not entail that it must respond in some characteristic way to that information. If there is a metacontrol mechanism, as some theorists have suggested, then further study may well reveal new truths about it that could explain why in some situations it engages several systems in response to a set of stimuli and in others it only engages a single system in response. Moreover, it should not be very surprising that this metacontrol mechanism does not initiate dual responses very often when conflicting stimuli are fed to the two hemispheres. Given the environment that such a control mechanism usually finds itself in, it would not be very surprising if it turned out to be relatively insensitive to certain kinds of stimuli like chimeric images. Since visual situations analogous to chimeric images seldom, if ever, occur outside of the experimental conditions split-brain subjects are placed in, it is not in the survival interest of organisms like us to develop sophisticated means of responding to such situations.

Objection 3 The previous question might be motivated by a more general worry: does this account not divorce consciousness from action and behavior in a way that is problematic? It maintains that a hemisphere can be conscious even when it remains completely “silent” in response to stimuli.

I am sympathetic to this worry. Any theory that attributes consciousness to a subject or system that remains unresponsive to stimuli and changes in the environment deserves to be treated with a degree of suspicion. However, the problem with respect to split-brain research, and even brain research in general, is that the results often do not respect the various sensibilities one might have with respect to the issue of the relationship between consciousness and behavior. For instance, the theorist who is hesitant to attribute consciousness to an individual or system that does not respond in appropriate ways to stimuli can adopt the Switch Model and rest easy with many of the findings from split-brain research. But then, as I have shown above, the results from the interrupted chimera trials seem to push such a theorist toward saying that sometimes behavior that is otherwise taken to be indicative of consciousness can be undertaken by systems that were unconscious at the time the stimulus they are responding to was shown. Conversely, the theorist who is sympathetic to the alternative view being defended here, can rest easy with the results of the interrupted chimera trials, but must maintain that, in many other situations, systems that are behaviorally silent are, nevertheless, conscious. Neither position is entirely comfortable. This being said, I favor the latter over the former. I think it does a better job of accommodating the data and, given that a coherent and plausible story can be told about the mechanism and systems responsible for governing behavior, the worries about attributing consciousness to behaviorally silent systems are muted.

4.4 Before and After the Experimental Conditions

All of the views discussed up to this point have operated under the assumption that, if anything changes the structure of consciousness for split-brain subjects, it is the surgery and not the experimental conditions. This assumption, however, is not universally held. Tye (2003), for instance, argues that it is the experimental conditions that change the structure of split-brain subjects' conscious experience and that the

changes revert once the experimental conditions are no longer in play. The metaphor he employs is that of stream or river that forks around an island or sand bar and comes together again on the other side. In Tye's opinion, the two most plausible views about split-brain subjects are (i) that the surgery splits a single stream of consciousness into two at the time of the surgery and that the split persists thereafter or (ii) that the stream of consciousness is only split during the experimental conditions and merges again after. Tye favours the latter. In his view, there is an extremely close relationship between experience and behaviour and he explicitly rejects a view like the one I have presented and defended above according to which the phenomenal consciousness and access-consciousness of split-brain subjects come apart. According to him, when there are cases in which it is clear that a subject lacks access-consciousness of some object, "access-consciousness is missing [in the subject], *because* phenomenal consciousness is missing too" (original emphasis) (2003, 125). Phenomenal consciousness, then, can be seen as a kind of gateway to access-consciousness. So if, as appears to be the case with split-brain subjects, one hemisphere appears to have access to some sensory information that the other does not, there is reason to believe that there is a corresponding experience generated by the one hemisphere and not by the other. Thus, if split-brain subjects display disunified behaviour, there is reason to believe that there is a disunity of experience. And if their overall behaviour is unified, we have *prima facie* reason to believe that their overall experience is unified.

Tye anticipates the objection that his view does not seem to account for the fact that it is the surgery and not the experimental conditions that brings about a permanent change to the organizational structure of the brain. The surgery seems to make the two hemispheres more causally isolated from each other, so, if one is prepared to entertain the view that split-brain subjects have two streams of experience, the most natural view to hold would be the view according to which a split occurs at the surgery and persists after. In response to this worry, Tye offers a certain picture about

the relationship of the two hemispheres to experience. He claims that the two hemispheres each generate experience, but, because their respective experiential products are redundant—outside of the experimental conditions each of the two hemispheres produce qualitatively identical experiences—only one experience is produced. He suggests that the two hemispheres working together to produce a single experience can be thought of as analogous to two synchronized movie projectors showing the same movie on the same screen; although there are two projectors, there is only one image. The implication with respect to Tye’s view of split-brain subjects is that it is only in the experimental conditions that the experiences produced by the two hemispheres are no longer redundant. When this happens, the single experience becomes two.

I find Tye’s view to be problematic. His view, it seems to me, is too reliant on observable behaviour and not sensitive enough to what we know about the nervous system. According to Tye, we can be confident that, during periods of observable behavioural disunity, the experiences produced by the two hemispheres are not redundant. Presumably, the disunity is due to a divergence in the sets of sensory information processed within the respective hemispheres. But what should we say about the sets of sensory information that are processed in the respective hemispheres outside of the experimental conditions when the subjects exhibit unified behaviour? Should we say that they are identical and, therefore, that the experiences generated by the respective hemispheres are identical and redundant? We should not. We know, for instance, that there is no time at which the two hemispheres of a split-brain subject receive the very same visual information. During normal conditions, the two hemispheres can make up for visual deficits via eye saccades. Thus, the two hemispheres are not under as much pressure to produce disunified sets of behaviour. However, at any given time, the respective sets of visual information will be as different from each other as they are during the experimental conditions. It is just that the experimental conditions are set up so as to compel the subjects to act on diver-

gent sets of visual information. Thus, even if we accept the kind of view that Tye advocates, one according to which the effect that redundancies in experience cancel out to produce only a single overall experience, a view I have reservations about, we do not have reason to believe that there are the perfect redundancies outside of the experimental conditions that Tye seems to believe there are.

Tye has not given us good reason to believe that the experimental conditions change the structure of conscious experience for split-brain subjects, and I am in agreement with most theorists in maintaining, contra Tye, that the experimental conditions do not themselves change the structure of conscious experience. Where I part company with most theorists, however, is over our views regarding what the structure of consciousness for split-brain subjects is generally like. They maintain that the conscious experiences of split-brain subjects are phenomenally disunified in certain ways, whereas I maintain that they are phenomenally unified.

CHAPTER 5

HOTS, REPRESENTATIONAL MISMATCHES, AND THE UNITY OF CONSCIOUSNESS

In the last two chapters I made a case for the Unity Thesis. I argued that interpretations of the split-brain data that are inconsistent with the thesis are problematic and that there is a way of interpreting the data that is consistent with it. Consciousness is a rich and multi-faceted phenomenon. Even if the Unity Thesis is true, there is still much for a complete theory of consciousness to address. One might, for instance, want an account of *why* consciousness is unified. Few have attempted to construct theories of consciousness that address more than a few of its many features. One theorist who has done more than many in this regard is Rosenthal. Rosenthal is, perhaps, most well known for his higher-order thought (HOT) theory of consciousness, a theory whose primary objective is to address the question of what it is that makes a mental state a conscious one, but he has also recently extended his theory of consciousness to account for its apparent unity. In this chapter, I will consider Rosenthal's account of conscious unity. Although there are questions one might have about the account in terms of whether it describes a plausible mechanism whereby states of consciousness are unified, this aspect of his account of conscious unity will not be the primary target of my arguments. Instead, my aim is to consider his theory of conscious unity in light of a well-known objection to his general theory of consciousness, an objection I will call the "Representational-Mismatch Objection" or, for the sake of convenience, the "Mismatch Objection."

There is a considerable literature on this objection. Some, myself included, are of the opinion that the objection is fatal to higher-order theories of consciousness.

Higher-order theorists disagree. I will present the Mismatch Objection and then weigh in on some of the most recent attempts to defend higher-order theories from it. Of particular interest will be Rosenthal's own responses to the objection. I will argue that these responses on behalf of higher-order theories fall short. This is significant, because, as I will argue, some of the mental mechanisms that Rosenthal invokes to account for the unity of consciousness generate representational mismatches. If the Mismatch Objection cannot be satisfactorily discharged, then some aspects of Rosenthal's account of the unity of consciousness make his theory of consciousness particularly subject to the objection. Although my criticisms will be aimed at Rosenthal's account, they will have somewhat broader implication. The broader implication is that, even if the mental mechanisms posited by Rosenthal's account of conscious unity ultimately show promise with respect to their ability to account for the unity of consciousness, some of them may be off limits to higher-order theorists.

5.1 The Mismatch Objection

Higher-order theories of consciousness make consciousness a representational matter; a (lower-order) mental state gets to be a conscious state in virtue of there being a (higher-order) mental state that represents it in the appropriate way. According to some theories (see Armstrong, 1968, 1981; Lycan, 1987, 1996), the higher-order state must be perception-like in nature. According to Rosenthal's theory (1986; 1997; 2005a), the higher-order state must be a thought about the lower-order state it represents.¹ Although higher-order perception and higher-order thought theories disagree over the nature of the higher-order state, they are both two-state theories in that

¹Carruthers (1996, 2000) has also defended a higher-order theory according to which the higher-order state is thought-like in nature. The difference between his theory and Rosenthal's, however, is that, whereas Rosenthal maintains that, in order for a state to be conscious, it must *actually* be accompanied by a HOT, Carruthers argues that all that is required for a state to be conscious is for one to be *disposed* to have a HOT about it.

they agree that the higher-order, representing state is numerically distinct from the lower-order state it represents. One-state higher-order theories (see Kriegel, 2009) part ways with two-state theories and maintain that the lower-order, represented state and the higher-order, representing state are one and the same²; according to one-state theories, a state is conscious if and only if it represents *itself* in the right way.

The Mismatch Objection has been pressed against higher-order theories by numerous people (see Byrne, 1997; Neander, 1998; Levine, 2001; Block, 2011). Where representations are involved, the question can be asked whether it is possible for the representational vehicle to misrepresent its target. Since there does not seem to be any reason, *a priori*, to suppose that mental states are immune from representational error, the possibility is a live one. If it is possible for a mental state to misrepresent another, the question can be asked what it is like for the subject when there is such a representational disconnect between a higher-order state and the lower-order state it purports to represent. In particular, the question can be asked whether the phenomenal character of the subject’s experience conforms more to the representational content of the higher-order, representing state or the lower-order, represented state. Asking this question raises a problem for the higher-order theorist because it seems as though, no matter how the question is answered, the answer undermines the theory.

Suppose my visual system is working correctly and I look at my red coffee mug under normal lighting conditions. This will generate a first-order mental state in

²There is some deviation between theorists and commentators over how the terms “higher-order” and “lower-order” are used. Some use the terms strictly to signify a mental state’s ranking in the hierarchy of mental states. On this usage, it would be incorrect to call one-state theories “higher-order” theories, because no numerically distinct mental state from a higher rank in the hierarchy is required to make a mental state conscious. On this usage, it would be more appropriate, as some do (Block, 2011), to refer to one-state theories as “same-order” theories. According to another usage, the one I will use in this paper, the terms “higher-order” and “lower-order” are used to refer to mental states via their representational properties. On this usage, it is not incorrect to refer to one-state theories as “higher-order” theories, because, according to one-state theories, conscious mental states represent themselves and, thus, encode both lower- and higher-order representations.

me that represents my coffee mug as being red. Now suppose my first-order state is targeted by a higher-order state, but, instead of representing my first-order state as a state that represents my coffee mug as red, my higher-order state misrepresents my first-order state as a state that represents my coffee mug as green. What will the phenomenal character of my experience be? Will my experience have the property of phenomenal redness or phenomenal green-ness? If my experience has the property of phenomenal redness, then it becomes unclear what the role of the higher-order state is in making the first-order state conscious. For if the phenomenal character of my experience conforms to the representational content of my first-order state, we are left to wonder whether I could have had the very same experience if the first-order state had occurred in the absence of the higher-order state, and then the door is opened to the possibility that first-order states do not need to be targeted by higher-order states in order to become conscious. Suppose, instead, that my experience has the property of phenomenal green-ness, conforming to the representational content of my higher-order state. It would then seem as though my higher-order state is conscious as opposed to my first-order state. Since higher-order states are not, in general, themselves targeted by further higher-order states, the door is again opened to the possibility that a mental state can be conscious without being the target of a representing state.

Although the representational disconnect in the situation just considered is enough to reveal the seriousness of the problem for higher-order theories, situations in which the representational disconnect is even more drastic present the problem in even starker relief. In the situation just considered, even though the higher-order state misrepresents its target state, there is a target state to be misrepresented. But it also seems possible that a higher-order state could represent that some lower-order state exists when, in fact, there is no such first-order state. And just as before, the question can be asked regarding what it is like for the subject to be in such a higher-order state.

If the higher-order theorist maintains that there is nothing it is like for the subject in this situation, then it seems as though the higher-order state is irrelevant with respect to consciousness. If the higher-order theorist maintains that the phenomenal character of the subject's experience conforms to the representational content of the higher-order state, then, because higher-order states are not generally themselves the target of higher-order states, it seems as though higher-order representation is not necessary for consciousness.

It bears mentioning that there is some reason to believe that one-state theories fare better against this objection than two-state versions. I am not entirely convinced that one-state theories ultimately escape this objection, but I will not defend this conviction here. My primary target is Rosenthal's higher-order theory and I will, on occasion, suggest that the conclusions I draw have broader implications. I do not claim, however, that the broader implications extend so far as to include one-state theories.

5.2 Responses to the Mismatch Objection

There are two distinguishable lines of defence that a higher-order theorist might pursue. The first is to take on the Mismatch Objection directly and argue that the kinds of representational mismatches considered in the objection can be accommodated by higher-order theories. The second is, perhaps, a kind of fallback plan. Failing a successful response of the first sort, one could argue that representational mismatches never occur or occur so infrequently that higher-order theorists need not worry about them. I will say at the outset, however, that if no successful defence of the first sort can be given, taking cover behind a defence of the second sort seems to me to amount to little more than whistling loudly while looking the other way. What is at stake in the debate over the Mismatch Objection is what is responsible for conscious experience. Even if higher-order mental states are *actually* immune to representa-

tional error—something few theorists, even higher-order theorists, believe—merely possible situations involving representational mismatches constitute, in my opinion, legitimate test cases. Higher-order theories purport to say something about what is responsible for conscious experience and so the theories must have something to say about what would be the case in the event that a representational mismatch between a higher- and lower-order state were to occur. That said, theory adoption often involves a messy cost-to-benefit analysis and, if there are no clear winners among the contenders, a defence of the second sort might constitute a way of mitigating the damage wrought by the Mismatch Objection and go some distance to helping higher-order theories in the cost-to-benefit battle. I will consider three direct responses and then I will consider a defence of the fallback variety.

5.2.1 Direct Defences

5.2.1.1 Rosenthal's Response

The first response to the Mismatch Objection I will consider is one given by Rosenthal. Although Rosenthal offers this response in defence of his HOT theory of consciousness, the response could be appropriated, with very little modification, by other higher-order theorists. Rosenthal maintains that, in situations in which a HOT misrepresents its target state, the HOT has the final say on what it is like for the subject (2005b, 213). Regardless of what particular first-order state one happens to be in at a time, what it is like to be in that state is determined by how one is aware of that state, and how one is aware of that state will be determined by how one's HOT represents that state. Dental patients, he points out, sometimes mistake the combination of the vibration of the dentist's drill together with their own fear as pain. They do this even though local anaesthetic has been properly applied. But it can happen that, after a patient is told that it is impossible for them to be feeling the nerve damage caused by the drill, the vibrations and fear are no longer experienced

as pain. The hypothesis is that, although the first-order states of the patient might be the same before and after being told about the anaesthetic and about how they might be misinterpreting their sensations, there is a difference between the way in which their HOTS represent those first-order states.

According to Rosenthal (2005b, 217), maintaining that HOTS have the final word on what it is like for the subject does not, contra Neander (1998) and Levine (2001), show that first-order states are irrelevant with respect to an experiential state's qualitative character. The reason for this is that those who press the Mismatch Objection presume that qualitative character must be conscious. They presume that qualitative character and what it is like for the subject are the same thing, and this fails to take into account that, according to the HOT theory, qualitative character and what it is like for the subject can diverge. First-order states, not HOTS, encode qualitative character. They are the ones that have qualitative properties. But not all sensory states, and thus not all states with qualitative properties, are conscious. Thus, it would be a mistake to presume that qualitative character and what it is like for the subject are perfectly correlated. What it is like for the subject will be determined by how the subject is aware of being in that particular first-order state. A subject can be completely unaware of a particular first-order state, a subject can misrepresent it in some way, or, in more extreme cases, it might seem to the subject as though he or she is in a state that he or she is not actually in. In each of these cases, what it is like for the subject will be different.

This response, it seems to me, misses the point. It may well be the case that those who press the Mismatch Objection take qualitative character to be conscious essentially, but few, if any, would deny that it is possible for an unconscious mental state to represent an external object's qualitative properties. The issue at hand is not whether it is possible for an unconscious mental state to encode sensory information. As I understand him, Levine uses the expression "qualitative character" to refer to

the character or quality of an experience. It refers to the way in which an *experience* of a red object differs from an *experience* of a green object. One experience has what Levine calls “reddishness,” the other “greenishness,” both of which he takes to refer to experiential properties, properties that are “way[s] things appear to be” (2001, 108). Thus, the challenge to the higher-order theorist has nothing to do with states that do not appear to the subject to be any particular way or with states, to use Rosenthal’s terminology, that the subject is unaware of. The issue is about qualitative properties that are phenomenologically manifest. It is about phenomenal character or what it is like for the subject. Whatever it is that people use these different expressions to refer to, the moral of the Mismatch Objection is that higher-order theories of consciousness fail to give a relational account of *it*. Thus, when Rosenthal maintains that HOTS determine what it is like for the subject, it sounds to me like an admission of what those who press the Mismatch Objection are arguing for: if HOTS determine what it is like for the subject, then a state does not have to be represented by another, higher-order state to generate experience.

Rosenthal has responded in a similar fashion to a more recent pressing of the Mismatch Objection by Block (2011). Block considers a situation consisting of exactly one mental state: a HOT that represents its subject as seeing a green object. Although the HOT represents its subject as being in a lower-order state that represents a green object, the lower-order state does not exist. Given Rosenthal’s assertions that HOTS have the final say on what it is like for the subject, the existence of this HOT would seem to be sufficient to produce what Block calls a “conscious episode” (2011, 424). But according to Rosenthal’s HOT theory, a conscious episode requires there to be a lower-order state that is represented by a HOT, and this contradicts the supposition we started with: that this situation consists of only one mental state, not two.

Rosenthal (2011) responds by arguing that, although the existence of a HOT is sufficient for there being something it is like for the subject, the existence of a HOT

does not entail that it has a target (i.e., that the state it purports to represent exists) or that it accurately represents its target state (if the target state exists). According to Rosenthal, the reason Block believes an episode of what-it's-like-ness commits the HOT theorist to the existence of a first-order state is that Block thinks of what-it's-like-ness in the same way he thinks of phenomenal consciousness (Block, 1995), namely, “as a monadic property that attaches to a qualitative state” (Rosenthal, 2011, 434). And if we make the further assumption that all qualitative states have this property, then we get the result that an episode of what-it's-like-ness must include a first-order qualitative state. Rosenthal then makes a point similar to the one he made in response to Levine: phenomenal consciousness is nothing more than the occurrence of mental qualities.

But states with mental qualities sometimes occur subliminally, that is, when one subjectively takes oneself not to be in any such state. And it's quixotic to regard as conscious a state that one subjectively takes oneself not to be in. Since phenomenal consciousness consists in qualitative character independently of there being any way the state seems to one subjectively, the same holds for Block's what-it-is-like-ness. (Rosenthal, 2011, 434)

Thus, according to Rosenthal, Block has not shown that the existence of an episode of what-it's-like-ness entails the existence of a (first-order) qualitative state.

Again, I find Rosenthal's response to be problematic. The focus on what Block's own views regarding what-it's-like-ness may or may not be is somewhat of a red herring.³ The objection does not rely on those views. The property of mental states that the expression “what-it's-like-ness” is used to refer to is a property that we all

³Incidentally, I doubt Block would assent to Rosenthal's assertion that “Block's phenomenal consciousness is itself arguably just the occurrence of mental qualities, independently of whether those mental qualities occur consciously” (Rosenthal, 2011, 434).

have a pre-theoretic grasp of. It is just the distinctive feature of our inner mental lives that arrives as we wake in the morning and disappears as we fall into dreamless sleep. It has been called many different things. “What-it’s-like-ness” is but one expression that has been used to refer to it. How this distinctive feature is to be accounted for within a theory of consciousness is what is up for grabs. The question is whether a mental event, such as the one described by Block, that consists of only one HOT has this distinctive feature. If, as Rosenthal seems to suggest, it does, then clearly it does not require two mental states being in a certain representational relation to each other to produce it, and the HOT theory of consciousness is undermined.⁴ If it does not, then Rosenthal owes us some account of what he means when he says there is something it is like for the subject of such a HOT, for he is then using the expression in a way that differs significantly from the way it is commonly used in the literature.

5.2.1.2 Weisberg’s Response

The second response I will consider is one recently offered by Weisberg (2011). Like Rosenthal, Weisberg is unmoved by Block’s version of the Mismatch Objection. Weisberg takes higher-order theories to be giving an account of what-it’s-like-ness, where the expression is given, what he calls, a “moderate” reading. According to this understanding, the expression is just a “common-sense way of picking out conscious experience” (2011, 439). Thus, Weisberg gives the impression that higher-order theories of consciousness are an attempt to give an account of the distinctive feature of mental states I discussed at the end of the previous section. He also claims, along with Rosenthal, that what it is like for the subject is a matter of how a subject’s higher-order states represent his or her lower-order states, and, in a case of misrep-

⁴Of course, a HOT theorist could deny that the HOT theory of consciousness is a theory about this distinctive feature of our mental lives, but then the theorist would owe us an account of what the theory *is* an account of. Moreover, the burden to do this would be significant, since so many have taken higher-order theories to be giving an account of what-it’s-like-ness as it is conventionally understood.

resentation, what it is like for a subject is how he or she “(erroneously) represents things as being” (2011, 440). So if a higher-order state represents one as being in a certain first-order state that does not exist, there will be something it is like for the subject and what it is like for the subject will be determined by the representational content of the higher-order state.

The following question can be asked: in a situation in which a higher-order state represents its subject as being in a state that does not, in fact, exist, which is the conscious state? Given that there is no lower-order state and that the higher-order state is responsible for what it is like for the subject, it would seem that the higher-order state is the conscious state. According to Weisberg, however, the answer is not quite so straightforward. He asks his readers to consider the following hypothetical situation. It is some time in the future when neuroscientists have a much fuller knowledge of the relationship between brain states and mental states. They have discovered, for instance, that it is possible for a subject’s visual system to be in a certain state without there being something it is like for the subject and that, when there *is* something it is like for the subject to be in that state, that state of the visual system is always accompanied by a certain higher-level brain state. These neuroscientists decide to bring about the higher-level brain state independently of the state in the visual system. To the subject, it seems exactly as it does when that higher-level brain state is accompanied by the relevant state in the visual system. “Indeed, oscillating between the presence and absence of the [state of the visual system] causes no *subjective* difference” (original emphasis) (2011, 442).

Suppose the neuroscientists now ask which of the states in this situation is conscious. Weisberg suggests that it does not really matter what answer is given. He says there is a sense of “conscious state” according to which it would be appropriate to say that the higher-order state is conscious, since it is the one responsible for what it is like for the subject. However, he says, there is also a sense of “conscious state”

according to which it is appropriate to say that the lower-order state is conscious. “Even though this state does not exist, it is conscious because the subject represents herself as being in that state” (2011, 442).

Suppose we grant Weisberg his sense of “conscious state” according to which the lower-order state in this situation can be said to be conscious. On this usage, all it means to say of a first-order state that it is conscious is that it is the representational object of a state that purports to represent it, a property it can have whether or not it exists. Making sense of reference to objects that do not exist has a long and tortured history in the philosophical canon, and whatever one might be inclined to say about such reference and the objects of such reference will certainly apply here. However, to be the object of a content-bearing entity is, to be sure, no more problematic for a mental state than any other kind of object, and it is certainly no more problematic for a mental state to purport to refer to a non-existent object than it is for a linguistic entity to do so. What is problematic, however, is that this sense of “conscious state” has little discernible relevance to the discussion at hand. Non-existent mental states cannot in any way be responsible for the fact that there is something it is like for a subject nor for the particular character of that what-it’s-like-ness. Mental states that are conscious in this sense, at least the non-existent ones, have no work to do in a theory of consciousness that purports to give an account of phenomenal character or experience.

There is another way one might understand the kind of statement that Weisberg makes in the above quotation. It is suggested by Matey (2006) in a discussion of Rosenthal’s HOT theory. Matey suggests that Rosenthal’s theory is ambiguous. One can understand Rosenthal’s theory as a theory about what makes mental state *tokens* conscious or about what makes mental state *types* conscious. The latter understanding has the advantage of avoiding the above questions about situations in which no lower-order state exists to be the relevant target of the higher-order state. If what is

made conscious is the content of a type of mental state rather than the content of a non-existent lower-order state, then there is no need to say that a non-existent token state is conscious in virtue of a higher-order representation. Although it stretches the grammar to read Weisberg’s statement in this way, and Weisberg does not himself explicitly endorse Matey’s suggested understanding, if read in this way, Weisberg’s statement would not be subject to the kind of question I raised for it above. This is not to say, however, that reading Weisberg’s statement in this way or extending this interpretation to higher-order theories more generally answers all the questions one might have. Even on this understanding, says Matey, higher-order states “purport to represent content of other states” and are themselves “special kinds of states capable of bearing phenomenal properties” (2006, 168). And what the model suggested by this understanding leaves unexplained is “how the token higher-order mental states that bear those contents acquire their phenomenal properties; it does not explain why or how those token states became conscious states” (ibid.).

5.2.1.3 Representational Mismatch as Partial Representation

In this section, I will consider another way of attempting to diffuse the Mismatch Objection. It involves attempting to cast cases of putative representational mismatches as *partial* representations. Cases in which a higher-order state partially represents a lower-order state can be understood in a way that does not pose a problem for higher-order theories. Thus, if all representational mismatches can be understood as cases of partial representation, then the Mismatch Objection can be diffused. But not all cases of representational mismatches can be understood in this way, most notably, the cases that are typically raised in discussions of the Mismatch Objection. Thus, this strategy for defending higher-order theories from the Mismatch Objection ultimately fails. I will, however, take some time to discuss this strategy for it will be relevant to my later discussion of Rosenthal’s theory of the unity of consciousness. In

section 5.3, I will argue that Rosenthal's account of conscious unity invokes mental mechanisms that generate representational mismatches, and that some of these representational mismatches raise the same questions as the cases that are commonly presented in discussions of the Mismatch Objection. Thus, it will be worthwhile to spend some time discussing partial representations in greater detail here. The possibility of a HOT partially representing its target state is explicitly endorsed by Rosenthal (1993, 2005b) and a version of this response is discussed briefly by Byrne (1997). I will develop it in a slightly different way.

Rosenthal (2005b, 211) maintains that HOTS often play an interpretive role; they are responsible for how one is aware of one's lower-order states. One of the ways in which they play this interpretive role is by selecting which lower-order content to represent at the higher-order level. HOTS can do this in two ways: they can select which lower-order states to represent, leaving some of them unrepresented at the higher-order level, or, of any particular lower-order state they select, they can opt to represent only part of the content that it encodes.

Rosenthal offers some examples of this mechanism at work. Consider the novice wine taster. After acquiring the concept of a tannin, his experience of a particular wine is likely to change. He is more likely to notice the tannins in the wine. This is not because his first-order, sensory states represent the qualities of the wine differently. It is because, according to Rosenthal (2005b, 187), he has another concept that can be invoked when his HOTS represent his first-order states; his HOTS can represent his first-order states in a different way. Prior to acquiring the concept, his HOTS are unable to represent his first-order states as representing the tannins in the wine. After acquiring the concept, they are. Rosenthal also considers what it is like for the concert goer after she acquires the concept of an oboe. Prior to acquiring the concept, she may be unable to pick out the oboes from the orchestral palette, but after acquiring the concept, she may well be able to do so. Again, this change in

what it is like for her is not due to a difference in the way her first-order, sensory states represent the sounds the orchestra makes. Rather, the change is due to some previously unrepresented content of her first-order states coming to be represented at the higher-order level.

Selective or partial representation does not violate the model proposed by higher-order theories of consciousness. The Mismatch Objection derives its force from the idea that, if the higher-order state is seen to be responsible for what it is like for the subject, then the lower-order state seems irrelevant. However, in cases of partial representation, it can be argued that the higher-order state is not entirely responsible for what it is like for the subject. For then, even though what it is like for the subject corresponds more perfectly to the content of the higher-order state, every bit of content encoded by the higher-order state appears at the lower-order level. The higher-order state does not encode any *new* content. It merely encodes a proper subset of the content that is already encoded at the lower-order level. In this way, the representational basis for what it is like for the subject can be seen to rest, ultimately, at the lower-order level, and the higher-order theorist can maintain that what it is like for the subject depends ultimately on what content the subject's first-order states encode. The higher-order states can be seen as playing a merely cooperative role, not as being solely responsible for generating experience.⁵

The cases of representational mismatch that this strategy is in the best position to handle are those in which it is plausible to suppose that higher-order states do not represent more than is already represented by lower-order states. The cases of the wine taster and the concert goer are two examples of this. Another set of cases that can be thought of in this way are those drawn from empirical studies on inattentional

⁵Admittedly, when Rosenthal talks about the role of HOTs within his theory of consciousness he talks about how HOTs “determine” (2005b, 187) and how they have “the last word” (2005b, 213) on what it is like for the subject. Thus, if he were to employ this kind of defence against the Mismatch Objection, this way of speaking could not be taken at face value.

blindness. In these studies, subjects are presented with certain objects or stimuli in plain view, yet, because the subjects are focused on a particular task or a different set of objects or stimuli, they demonstrate a lack of awareness of them.⁶

Although these cases of representational mismatches between lower- and higher-order mental states can plausibly be understood as cases of partial representation, there are still questions that can be raised. Should it turn out that, in some of these cases, some unrepresented first-order content is, nevertheless, phenomenologically present in the subject's experience, the higher-order theorist would again seem to be in a position of having to admit that it is possible for a mental state (or parts of a state) to be conscious without being represented at a higher-order level.

Subjects in the inattentional blindness studies profess to be unaware of (or can be shown to be unaware of) objects that appear in the centre of their visual field. Since they profess complete unawareness, it is difficult to know for certain whether or not the objects they profess to be unaware of have any presence in experience. Nevertheless, one can argue that these objects have some presence in experience even if they fail to have a presence in certain higher-order or downstream mental states. Consider an analogous case: Armstrong's (1981) long-distance truck driver who "comes to" and realizes he does not remember anything about the last several miles. Even though the driver may be unable to recall features of the road and surrounding landscape or remember stopping at a traffic light, he or she successfully negotiates the road. It is hard to believe how this would be possible unless the driver is conscious (in some sense of the term) of the features of the external environment needed to keep the truck on the road (Dretske, 1993). In the same way, it is difficult to believe that, in typical inattentional blindness studies, the objects displayed in plain view to the subjects are not phenomenologically manifest in any way whatsoever.

⁶For some examples of these studies, consult Simons and Chabris (1999), Hyman Jr. et al. (2010), and Chabris et al. (2011).

Rosenthal (2002) considers a similar case. While walking through the woods, one may step over numerous branches without having any recollection of them, especially if one is deeply involved in a conversation with a hiking partner at the time. What is clear is that one must be in the appropriate sensory states to avoid tripping over the branches. The question is what the nature of one’s experience is like at the time. Rosenthal suggests that one has a visual experience of the environment—there are no gaps in one’s visual field where the branches lie—but that the experience is just of an “undifferentiated rustic environment” (2002, 416). What might be responsible for the difference between one’s visual states and one’s experience? One’s HOTS select which elements of one’s sensory states to represent.

Cases like that of the wine taster and concert goer and cases taken from standard inattentive blindness studies are the most amenable to the strategy of accounting for representational mismatches in terms of partial representation. But other kinds of representational mismatches, namely those that are typically cited when the Mismatch Objection is raised, are much less amenable to this strategy. It is entirely unclear, for instance, how partially representing a red-representing state could result in a higher-order state that represents it as being green-representing. The limits of the partial-representation strategy are even more obvious when considering representational mismatches of the most extreme kind: situations in which the lower-order state does not even exist. Thus, although some cases of representational mismatch can be described as cases of partial representation, not all of them can.

5.2.2 A Fallback Defence

The three responses to the Mismatch Objection I have considered are direct responses in that they attempt to show that higher-order theories can accommodate the kind of representational mismatches that the Mismatch Objection exploits. In this section, I want to consider a different line of response to the Mismatch Objection,

one that a higher-order theorist might fall back on if the kind of direct responses I have been considering fail. I have already drawn attention to the limitations of this response, but there are two reasons for considering it here. First, as I indicated, the business of deciding which theory to accept often involves weighing a multitude of pros and cons. Thus, any response that goes some way toward addressing an objection to a theory is worth considering. Second, discussion of this fallback response will be relevant to my discussion, below, of Rosenthal's theory of conscious unity.

The representational mismatches that are considered when the Mismatch Objection is raised are typically rather significant ones. If it could be successfully argued that these representational mismatches are relatively rare, something which, given their magnitude, might not be all that implausible, then the higher-order theorist might have a way to live with the theory even if the direct defences fail. Although there are a number of different considerations one might offer to show that representational mismatches between lower- and higher-order mental states are relatively infrequent and, when they occur, often relatively insignificant, given my focus on Rosenthal's work in this paper, I will only present a few of his considerations here.

Rosenthal (2005b, 212) argues that there are mental mechanisms that work to guard against and correct large-scale representational errors. Mental states play off of each other and stand in certain causal relationship to each other. Pains, for instance, make it difficult to concentrate on other things and often cause us to desire their cessation. This means that, if a significant representational mismatch were to arise between a lower- and higher-order state, a certain amount of psychological tension would accompany the mismatch. This psychological tension, maintains Rosenthal, would exert some pressure upon the relevant mental states to fall more into line with each other. Thus, if a HOT were to misrepresent a first-order pain, the psychological discord this would generate would draw attention to the representational mismatch and, most likely, would have a corrective influence on the HOT.

Rosenthal (2005b, 218–9) also presents a reason to think that representational mismatches between HOTS and their target states will be relatively infrequent. The reason has to do with the concepts that are invoked by our HOTS. When I am in the presence of a red object (in normal conditions), my sensory state represents the colour of the object. It does so in virtue of having a certain property, say, red*. In order for me to have a HOT that characterizes me as being in a red* state, I must possess the concept of a red* state. This concept is relatively easy to acquire, argues Rosenthal. All it takes is for one to be aware of the difference between being in the presence of a red object and being in a state that one is normally in when in the presence of a red object, and this awareness is naturally derived after being in a few situations in which one is in error about being in the presence of a red object. Because concepts such as the concept of being in a red* state are easy to come by and because of the natural command we have over the concept’s application, being in a red* state is usually accompanied by a HOT to the effect that one is in a red* state. Thus, there will seldom be a representational mismatch between our lower- and higher-order states.

I have several comments. First, it is not clear to me that concepts regarding being in certain first-order sensory states are as easy to come by as Rosenthal makes them out to be. A well-known objection to higher-order theories of consciousness, especially HOT theories, is that, because they seem to make certain intellectual capacities and certain concepts a requirement for conscious experience, they have the result that most, if not all, non-human animals and young children are not capable of conscious experience.⁷ If a subject must, for instance, have the concept of a mental state in order to form the appropriate HOT needed to make a lower-order state conscious, then it would seem that the subject must have something like (what developmental psychologists call) a “theory of mind.” In normal human subjects, a theory of mind

⁷See Seager (2004) for a relatively recent discussion of this objection.

arrives on the developmental scene at three to four years of age.⁸ This is a problem for HOT theories, but it seems to me that the comments Rosenthal makes about how an arbitrary concept of being in a certain sensory state is acquired may show that the concepts needed to form the required HOTs about lower-order states may arrive on the developmental scene even later than the age of three or four.

In a typical theory of mind test called the “false-belief task” a child is shown the location of some object while in the presence of another subject. The subject leaves the room and the object is moved to another location. The child is then asked where the absent subject will look for the object when he or she returns. The answer the child gives indicates whether the child is aware of the fact that the absent subject is in a different epistemic position than he or she is. Obviously, having such an awareness requires a certain level of conceptual and cognitive sophistication. My contention is that the kind of concept Rosenthal suggests a subject invokes when, for instance, he or she has a HOT to the effect that he or she is in a red* state is a concept that may arrive even later than the concepts that theory mind tests like the false-belief task test for.

Suppose that the story regarding how a young child acquires the concepts necessary to provide correct answers in false-belief tasks is similar to the story Rosenthal tells about how one acquires the concept of being in a red* state: one acquires these concepts by being in enough situations in which one’s expectations or beliefs turn out to be incorrect. In the case of theory-of-mind concepts, this account would require one to be in enough situations in which others act differently than one expects them to. After being in enough of these situations, the child learns that others may have different beliefs about the locations of certain objects than he or she does. But what kinds of situations must a child be in in order to acquire the concept of being in a

⁸See Wimmer and Perner (1983) and Gopnik and Astington (1988).

red* state? Rosenthal's suggestion is that the child must be in enough situations in which he or she is in error about the colour of an object. Note that this is not just a matter of being in situations in which the child is corrected about the *name* of a colour. The child must be in situations in which an object appears to him or her to be red when, in actuality, the object is some other colour. Situations like this, I submit, are much more rare than situations in which the child's expectations about the actions of others turn out to be incorrect.

Given the relative immediacy with which we make judgements about the colours of objects and the way in which we are hard-wired to see colours of objects as constant across many different lighting conditions, the circumstances in which the kind of colour-perception errors required for the acquisition of the concept of a red* state would seem to be relatively hard to come by. Few of us gain explicit knowledge about how we perceive colours as constant across different lighting conditions until we take a post-secondary course in philosophy or psychology. And we seldom tune in to the fact that our colour judgements can be incorrect until we judge some object to be a certain colour and then notice that it appears to be a different colour when seen in other circumstances. Noticing this change in how an object appears requires a certain amount of attentiveness and reflection, something that also arrives relatively late in a child's cognitive and psychological development. Thus, it seems clear that acquiring the concept of being in a particular sensory state, such as the concept of being in a red* state, would, on average, happen much later than the acquisition of the kinds of concepts theory-of-mind tests test for.

Second, although I find it entirely plausible that there are mental mechanisms that strive to keep different mental states in step with each other and minimize certain conflicts between them, I am doubtful that these mechanisms are likely to intervene and exert a corrective influence other than in the most extreme cases of representational mismatch. It is relatively easy, for instance, to see how a situation in which

a subject's first-order states represent something as being very pleasurable and the subject's HOTS represent those states as representing pains might cause a kind of psychological tension that would either bring the HOTS in line with the first-order states or vice versa. The vast majority of representational mismatches, however, will not produce the kind of psychological tension needed to exert the kind of corrective influence needed. Take the situation involving the dental patient that Rosenthal discusses. The patient's lower-order states represent the patient's felt fear and the vibrations of the dentist's drill. The patient's HOTS represent those states as representing pain. What psychological tensions will the representational mismatches between these states generate? It is not clear that they will generate any. The reason is that the causal roles of pain and fear are, in many ways, similar. Fear, just as pain, makes it difficult to concentrate and causes one to fixate on the object or stimulus that is its cause. Fear causes the desire for the object of the fear to disappear or cease to exist, and fear can cause one to avoid the fear-inducing object or stimulus in the future. None of these effects of fear are particularly incompatible or at odds with a state that represents one as having a pain or with a state that represents the vibrations of the dentist's drill. Thus, the subject, especially while in a state of fear, a state which makes sober reflection difficult, is unlikely to experience a tension in her sensory states and thoughts and, thus, it will be unlikely that the representational mismatch between her states will be resolved.

Rosenthal claims only that psychological tension will have a reconciling effect on relatively large representational mismatches and, in his discussion of the dental patient case, does not claim that the patient will experience corrective psychological tension. My point here is merely to make clear that there will be many situations involving representational mismatch in which this corrective mechanism will not be engaged. It is easy to come up with other examples of situations in which representational mismatch is unlikely to generate psychological tension. For instance, it is possible for

a hiker, expecting to arrive at a waterfall within minutes, to mistake the sound of traffic on a distant highway for the sound of a waterfall. And it is possible, having seen one's companion remove a tick from his or her arm, to feel ticks on one's own limbs even though there are not any there. In each of these cases, we can suppose that the subject's first-order states represent the relevant stimuli accurately—in the tick case, we can suppose that the subject's first-order states do not represent the crawling of an insect at all—and in each of these cases, new information can cause one to feel, hear, or experience things differently. When the dentist tells the patient that the nerves have been silenced via local anaesthetic, the patient may cease to experience the fear and drilling as pain. When told that there is a highway nearby, one might correctly hear the sound as traffic. And when one realizes there is no tick at the location of the sensation, the sensation may well disappear. In each of these cases, the revision in what it is like for the subject is not brought about by some psychological tension produced by the states involved in the representational mismatch themselves. The revision is brought about by new information after the mismatch has persisted for a while. Thus, although representational mismatches may invoke the relevant kind of correctional mechanism in relatively extreme cases, it does not seem to be the kind of mechanism that will, in general, correct representational mismatches that are of a less extreme kind.

5.3 Rosenthal's Theory of the Unity of Consciousness and the Mismatch Objection

I have argued that Rosenthal's and Weisberg's recent responses to the Mismatch Objection are unsuccessful. I have also argued that, although representational mismatches that can be understood as cases in which higher-order states partially represent lower-order states are not subject to the Mismatch Objection, it is far from clear that all representational mismatches can be understood in this way. The fallback

response, though perhaps significant in the final cost-to-benefit analysis, is not one that stands any chance of addressing the questions raised by the Mismatch Objection.

In the remaining pages of this chapter, I will bring some of these results to bear on Rosenthal's account of conscious unity. First, I will present Rosenthal's account of conscious unity. Rosenthal's account invokes two distinguishable mental mechanisms. I will present them in turn and then evaluate them. Both of them generate representational mismatches. I will argue that only one of them can be understood in a way that is immune to the Mismatch Objection. In spite of its immunity to the Mismatch Objection, however, it may still turn out to be inadequate as a means of accounting for conscious unity.

5.3.1 Rosenthal's Theory of Conscious Unity

Rosenthal sets out to provide an account of the appearance of, what he calls, "mental unity." As I pointed out in section 2.1, there are many different varieties of conscious unity. Rosenthal (2005c) calls attention to some that I did not introduce, but it will not be necessary to discuss them here. As best I can tell, Rosenthal uses this expression, "mental unity," as an umbrella term to refer to all the different varieties of conscious unity together. Thus, Rosenthal's account seems to be intended as a general account of the unity of consciousness, addressing many (or all) of the different varieties of conscious unity in one fell swoop. It is, of course, an open question whether the same mental mechanism can be responsible for each of the different varieties of conscious unity. However, the objections I will raise for Rosenthal's account will not hinge on whether it is adequate as an account of one or other particular species of conscious unity.

Finally, before presenting his account it is worth drawing attention to the fact that Rosenthal is explicit about the fact that he is merely offering an account of the fact that consciousness has the *appearance* of being unified. He leaves open the possibility

that our conscious states are not as unified as they appear to us to be; the unity of consciousness may be a mere illusion and our mental states may, in actual fact, not be as unified as our subjective sense of their unity would seem to indicate. I will, for the most part, ignore this subtlety. I will do so for two reasons. First, as Rosenthal himself points out, “it is arguable that the appearance of conscious unity is, itself, all the reality that matters” (2005c, 341). Second, what I have to say will apply just as readily to Rosenthal’s account as an account of the appearance or subjective sense of conscious unity as it would to Rosenthal’s account if it were an account of the actual unity of experience.

Rosenthal motivates the need for an account of the unity of consciousness by drawing attention to a certain feature of his theory of consciousness. At any given time during our wakeful hours, we are in numerous different conscious states. Every conscious state is accompanied by a HOT that targets it. The atomistic nature of this theory of consciousness, says Rosenthal, “may seem to prevent it from being able to explain our sense of the unity of consciousness. If each conscious state owes its consciousness to a distinct HOT, how could we come to have a sense of such unity?” (2005c, 340). How is it that “all of our conscious states seem to belong to a single, unifying self?” (ibid.).

Rosenthal answers this question by extending his theory of consciousness. Just as HOTs have an important role to play in making mental states conscious, they have an important role to play in generating our sense that consciousness is unified. HOTs, maintains Rosenthal, “operate on many of our mental states not singly, but in large bunches” (2000, 226). They target and represent many different mental states all at once. Rosenthal motivates this claim by considering the well-known cocktail party effect. At a cocktail party, where numerous conversations are going on at once, one tends only to hear them all as a background din. However, if one’s name is mentioned in a conversation across the room, one’s attention is immediately drawn to that con-

versation. Rosenthal suggests the following explanation. In order to hear one's name in that conversation, one must be hearing the articulated words of that conversation. Since one's attention would also be captured if one's name were to be mentioned in a different conversation, it would seem as though one hears the articulated words of many of the different conversations at the same time, albeit unconsciously. What one is conscious of, however, is a background din; the individual conversations are represented unconsciously. What this suggests, according to Rosenthal, is that "one's HOTS group many auditory sensations together, making them conscious only as an unarticulated bunch" (2000, 227). If this is right, then it is not the case that every one of one's perceptual states requires its own HOT to make it conscious. A single HOT can make a number of mental states conscious all at once.

Call this way of accounting for the unity of consciousness the "gathering strategy." It is a natural extension of Rosenthal's higher-order theory of consciousness. Rosenthal argues, however, that this strategy cannot account for conscious unity by itself. Here is what he says about the gathering strategy and its prospects as a complete account of conscious unity.

Wholesale operation of HOTS...doubtless helps to induce some conscious sense of unity among our mental states. But that will only go so far. Since no single HOT covers all our conscious states, the basic problem remains. How can we explain a sense of unity that encompasses states made conscious by distinct HOTS? (2005c, 342)

Even if we suppose that mental states can be unified in experience by being represented together by a HOT, if no single HOT represents *all* of our conscious mental states, then the gathering strategy will not provide a complete account of unity. For then there will always be some of one's mental states that have not been gathered together under a single HOT.

In response to this problem, Rosenthal presents another mechanism by which mental states come to be unified in experience. I will call this part of his account the “common-ascription strategy.” According to the common-ascription strategy, we come to experience mental states as unified in virtue of a feature of the contents of the HOTS that represent them and make them conscious. A HOT is a thought to the effect that its subject is in a certain mental state; it ascribes a mental state to an individual. The way it ascribes a mental state to an individual is by referring to its subject via something like the mental analogue of the first-person singular pronoun, “I.” In doing so, a HOT ascribes its target state to whomever it is that happens to be its subject.

There are two ways, according to Rosenthal, in which this feature of HOTS contributes to a sense of unity. First, the indexical manner with which a HOT refers to the bearer of its target state gives rise to the sense that one is directly aware of the bearer of that target state. Second, since a HOT ascribes its target state to its subject in an indexical manner, its content is “bare” with respect to its characterization of the subject of its target state. This has several consequences. In the same way that there is nothing about the sentence, “I am experiencing a pain in my right elbow,” that indicates, by itself, that it is about the same individual as the one referred to in another tokening of the same sentence, there is nothing in the content of one HOT that indicates, by itself, that the bearer of its target state is the very same individual as the bearer of another HOT’s target state. On the other hand, there is also nothing in the content of one HOT that indicates that the bearer of its target state is *different* than the bearer of another HOT’s target state. This, according to Rosenthal, facilitates a certain kind of subjective impression in us. “[O]ur seeming to be aware in a direct and unmediated way of the self each HOT refers to tilts things towards apparent unity. Since we seem to be directly aware of the self in each case, it seems

subjectively as though there is a single self to which all of one's HOTs refer, a single bearer for all our conscious states" (2005c, 344).

Thus, Rosenthal seems to offer a two-pronged account of the unity of consciousness: conscious unity is due in part to the fact that HOTs often represent a number of mental states all at once and in part to the fact that it seems to us as though all of our own mental states are had by the same subject.

5.3.2 The Gathering Strategy and Representational

Mismatches

The mental mechanism invoked by the gathering strategy generates representational mismatches. Consider again Rosenthal's account of the cocktail-party effect. Suppose you are engaged in conversation in the minutes leading up to the mention of your name across the room and that, just before your name is mentioned, there is a brief lull in the conversation you are having with your conversation partner. If Rosenthal's account of the cocktail-party effect is correct, then, barring significant differences in your respective senses of hearing, the first-order states of you and your conversation partner should be virtually identical; your respective first-order states both monitor the different conversations in the room. In the moments prior to the mention of your name, the HOTs in virtue of which you are both conscious of the conversations in the room should also be virtually identical, for you are both conscious of the conversations as a dull roar. However, when your name is mentioned, your HOTs no longer represent all of the conversations as a mere background din. Instead, you become conscious of the conversation in which your name is mentioned. The HOTs by which your conversation partner is conscious of the conversations in the room, however, may continue to represent those conversations as a dull roar. Because you are now conscious of a particular conversation and your conversation partner remains conscious only of a dull roar, there is a difference in phenomenal character between

your conscious mental state and your conversation partner's mental state. What makes for this difference in phenomenal character? Since your respective first-order states are virtually identical, it would seem as though the only place for this difference in phenomenal character to reside would be in the difference at the second-order level.

The mental mechanism employed by the gathering strategy is similar to the one employed in Rosenthal's account of the cocktail-party effect; according to the gathering strategy, the sense of conscious unity is due, at least in part, to the fact that HOTS can represent multiple first-order states at once. The implication is that, if the first-order states were not represented by a single HOT in this way, the sense of conscious unity would be lessened or eliminated altogether. But the sense of conscious unity that this mental mechanism gives rise to is phenomenologically real. Accounts of conscious unity are motivated, in part, by the assumption that things would have a different subjective feel if one's conscious mental states were not unified in the way they are. But where does this difference in phenomenal character lie? The answer, as it was in the discussion of the cocktail-party effect, is that the only place for the difference to reside is in differences at the level of second-order mental states.

In motivating the need for an account of the unity of consciousness, Rosenthal suggests that our sense of conscious unity seems to be at odds with the "atomistic character" (2005c, 340) of his theory of consciousness. If each conscious first-order state is made conscious by its own unique HOT, one might wonder why our experience is that of a unified global state rather than of a set of discrete states that are merely simultaneously conscious. When the gathering strategy is offered as an explanation for our sense of unity, the contrast, then, is between a situation in which each conscious first-order state has its own HOT and a situation in which multiple conscious first-order states are represented by the same HOT. But the difference between these contrasting situations is a second-order difference. In the former, each HOT targets only one first-order state. In the latter, a HOT may target many first-order states.

Moreover, this difference could persist even if the first-order states in the two situations were identical. Thus, the phenomenal difference that conscious unity makes, if it is accounted for with the gathering strategy, is due to a difference at the level of second-order mental states. It seems to make HOTS exclusively responsible for an aspect of what it is like for a subject. Since HOTS are not, in general, themselves the targets of higher-order states, it would seem as though this way of accounting for conscious unity opens Rosenthal to the charge that higher-order representation is not necessary for a mental state to be conscious.

In section 5.2, I considered several responses the higher-order theorist might make to the Mismatch Objection. It is worth pointing out that at least two of those lines of response are unavailable here. First, the contribution that the unity of consciousness makes to the phenomenal character of experience cannot be explained in terms of partial representation of first-order content or states. In the case of the novice wine taster or the concert goer, the fact that what it is like for the subject corresponds more perfectly to the representational content of the subject's HOTS can be argued to be unproblematic because, in each of these cases, the HOTS involved do not encode novel content, i.e., content that is not also encoded at the first-order level. All they do is select which first-order content to represent at the second-order level. With respect to the gathering strategy and conscious unity, however, the situation is significantly different; conscious unity, as it is accounted for by the gathering strategy, does not have a representational basis at the first-order level.

To see this, suppose the contrary. Suppose that the sense of conscious unity does have a representational basis at the first-order level. Then the sense of conscious unity could be made phenomenally manifest in the same way that phenomenal redness and phenomenal green-ness are. All that would be needed for the sense of conscious unity to become present in experience would be for some HOT (or HOTS) to represent the subject as being in some first-order states that have whatever property it is that

gives rise to the sense of conscious unity. There would be no need for first-order states to be represented *together* by a HOT to produce the sense of conscious unity. Since Rosenthal evidently believes that first-order states must be gathered together in order for a subject to enjoy a sense of conscious unity, it would seem that locating the representational basis for the sense of conscious unity at the first-order level in this way would undermine the need for the gathering strategy.

The second response that is not available to Rosenthal, here, is the response that seeks to argue that representational mismatches are relatively rare. In normal subjects, the unity of consciousness is pervasive and ubiquitous, so much so that some have argued that (for certain species of conscious unity, at least) consciousness is *necessarily* unified (Bayne and Chalmers, 2003; Bayne, 2010).⁹ Thus, an account of conscious unity must be an account that is compatible with the sense of conscious unity being present for most conscious subjects most of the time. If the gathering strategy is invoked in such an account, then any representational mismatches that accompany the operation of its corresponding mental mechanism cannot be rare or infrequent. Its corresponding mental mechanism does generate representational mismatches—at the lower-order level, representational content is represented in discrete bits by discrete states, whereas, at the higher-order level, those bits of content are conjoined with other bits of content into larger representational chunks—and so the fallback response is not available here.

There is, however, a way of understanding the gathering strategy and the way in which it accounts for conscious unity that may make it unproblematic from the point of view of the Mismatch Objection. As we saw above, the phenomenal difference that the unity of consciousness makes cannot be explained in terms of partial

⁹We have seen that this claim is not universally accepted. What is important here, however, is that many of the arguments against the necessary unity of consciousness are based on the experiential lives of non-normal subjects. The experiences of normal subjects appear (subjectively) to be unified and this is what Rosenthal is attempting to provide an account of.

representation of first-order content; there is no particular content at the first-order level that, when represented at the second-order level, produces the sense of unity. What Rosenthal can point to, however, is the fact that the mental mechanism invoked by the gathering strategy does not introduce any genuinely new content at the second-order level either. All the mechanism does, he could say, is represent first-order content together in a single second-order content. And it is the mere fact that first-order content is represented together that produces the sense of unity. Thus, there is a sense in which the phenomenal difference that conscious unity makes is due to the second-order representation, but there is no *particular* content, either at the first-order level or the second-order level, that is associated with the sense of unity. And thus there is no particular content that makes a phenomenal contribution to experience without itself being the target of a higher-order state.

There are two things to note about this response. First, it requires an assumption about the way in which phenomenal unity contributes to the overall phenomenal character of an experience. It is easy to think that phenomenal unity has its own special character, that, in addition to all the other phenomenal elements that contribute to the overall phenomenal character of an experience, there is the phenomenal character of phenomenal unity. The assumption that the above response requires, however, is that there is no such special phenomenal character associated with phenomenal unity. The reason for this is that, if Rosenthal countenanced such a special phenomenal character for conscious unity, he would be under some pressure to associate that special phenomenal character with a particular content of mental states. This would then saddle him with the challenge of attempting to integrate this content into his theory without making it one that could be conscious without being the target of a higher-order state.

This assumption, however, would not be a very onerous one for Rosenthal to embrace. Many theorists who have considered the unity of consciousness have pointed

out that it is problematic to understand the phenomenal contribution of phenomenal unity as being a separate or additional phenomenal character. The reason is that this conception of the phenomenal character of conscious unity would seem to invite a vicious regress.¹⁰ If there were such a special phenomenal character of phenomenal unity, it would also have to be unified with the rest of the phenomenal elements in experience. This would seem to require some sort of unifying relation to obtain between it and the rest of the phenomenal elements of one's experience, prompting the question whether this unifying relation itself has a special phenomenal character. It would seem that the answer would have to be affirmative, otherwise there would be nothing to distinguish between an experience in which the special phenomenal character of unity was unified with the rest of the phenomenal elements in that experience and an experience in which it was not unified with them. And once it is admitted that the unifying relation itself has a special phenomenal character, the regress is up and running.

The second thing to note in conjunction with this response is that, in spite of moving the gathering strategy out of the shadow of the Mismatch Objection, it may not be enough to save Rosenthal's account of conscious unity entirely. As I indicated in section 5.3.1, Rosenthal maintains that the gathering strategy is not able to account for conscious unity by itself. According to him, no single HOT is able to represent all of one's (conscious) first-order states at once. If this is true, then there will always be some first-order states that are not represented together in a single content at the second-order level. That is why Rosenthal partners the gathering strategy with the common-ascription strategy. There is some plausibility to the claim that HOTs are limited in the way Rosenthal suggests, but I will not take the time to examine it more deeply here. Here I will only note that if the common-ascription strategy, which

¹⁰This problem is discussed in section 2.2.3. See also Tye (2003, 21–25) and Siewert (2001, 547).

I will take up in the next section, generates representational mismatches that are problematic from the point of view of the Mismatch Objection, then the burden will ultimately rest upon the question of what the representational scope of higher-order states is.

5.3.3 The Common-Ascription Strategy and Representational Mismatches

The common-ascription strategy, you will recall, plays off of the way in which Rosenthal thinks of the contents of our HOTS. A HOT ascribes its target state(s) to a subject, and it does so by referring to the subject of its target state(s) indexically via something like the mental analogue of “I.” Rosenthal maintains that, in the same way that we are disposed to treat successive tokenings of the same name as referring to the same individual, we are disposed to treat successive occurrences of the mental “I” as referring to the same subject. Thus, we are disposed to take our HOTS as all referring to the same subject of mental states, and this gives rise to our sense of conscious unity.

There is a wrinkle that Rosenthal anticipates. “HOTS are not typically conscious thoughts; no HOT is conscious unless one has a third-order thought about it. So long as HOTS are not conscious, one will not be conscious of their seeming all to refer to a single self” (2005c, 344). Rosenthal suggests that the reason we have a sense of conscious unity even when we are not introspecting is that there is a kind of carry over from when we introspect. Within Rosenthal’s theory, introspecting is a matter of becoming aware that one is aware of one’s mental state. It is a matter of becoming the subject of a third-order thought-like state that targets a (second-order) thought one has about a first-order state. So, during introspection we become aware of the direct and indexical manner in which our HOTS refer to the individuals they ascribe their target first-order states to, and we come to have an “explicit sense of the unity of our conscious mental states” (ibid., 345). Because we have this explicit sense of

conscious unity in introspection, we develop an expectation that, whenever we are not actively engaged in introspection, we can readily become aware of that single bearer of our mental states in other contexts. This, says Rosenthal, leads to a dispositional and tacit sense of conscious unity even while we are not engaging in introspection.

Clearly Rosenthal conceives of the explicit sense of conscious unity as having a relatively pronounced phenomenological consequence. But this is unproblematic for Rosenthal, because the way in which active introspection is understood within the confines of his theory makes the phenomenal consequence of introspection perfectly explicable. What a subject is conscious of at any given time certainly impacts what it is like for the subject at that time, and introspection involves becoming conscious of some of our (previously unconscious) HOTs. Thus, accounting for the explicit sense of conscious unity in this way would seem to be immune to the Mismatch Objection.

Questions arise, however, when we consider the tacit sense of conscious unity. The tacit sense of conscious unity is what remains after the second-order states are no longer conscious, and it is associated with the expectation we have, when we are no longer actively introspecting, that we can, if we wish, become actively conscious of our mental states. The questions I want to raise about this tacit sense of unity have to do with Rosenthal's identification of it with a certain kind of expectation.

Some expectations have an identifiable phenomenal character. For instance, it is clear that children experience a certain set of feelings while they anticipate the imminent opening of a jack-in-the-box lid. Likewise, there is a certain set of feelings that accompany the expectation or anticipation of a reprimand by one's superior. But other expectations have a much more subtle phenomenal character if they have a phenomenal character at all. When I turn on my computer, I expect a certain sequence of events. It is unclear to me whether this expectation has any phenomenological impact. When the door of the elevator closes, I expect a certain sensation in the pit of my stomach as the elevator begins to move. The sensation in the pit of my

stomach has an identifiable phenomenal character, but it is not at all obvious that the expectation beforehand has any phenomenology.¹¹

Does the expectation that one can readily become aware of the bearer of one's mental states have a phenomenological impact or not? Suppose we say that it does. If it has a particular, identifiable phenomenal character, then Rosenthal's account faces two lines of objection. It would then seem as though he would be affirming that the *sense of unity* has a particular, identifiable phenomenal character, and this would open him up to questions about whether or not his account invites the kind of vicious regress that was discussed briefly in the previous section. Rosenthal would also be faced with the challenge of accounting for this special phenomenal character in a way that does not invite the Mismatch Objection. We are not conscious of our HOTs when we are not actively introspecting, and so the special, indexical, first-person content of HOTs that the common-ascription strategy relies upon is not available. If the expectation has a phenomenal presence or a phenomenological impact of some sort that is not a particular, identifiable phenomenal character, then Rosenthal would again be faced with the task of saying how the indexical first-person content of HOTs (that is not typically represented by third-order states or anywhere to be found in lower-order states) could be responsible for it. It should be noted that the situation here is importantly different than it was in the case of the gathering strategy. The gathering strategy accounted for the phenomenal impact of conscious unity by gathering the contents of first-order states into larger, conjunctive contents at the second-order level. This means that the gathering strategy accounts for the phenomenal manifestation of conscious unity via second-order representations. However, the difference between the gathering strategy and the common-ascription strategy is that the gathering strategy does not rely on any second-order content that is not also present at

¹¹This may be different for people with phobias of elevators or a with particular aversion to the feeling that the movement of the elevator causes.

the first-order level; the common-ascription, with its reliance on the indexical first-person content of HOTs, does. Finally, if we suppose that the expectation associated with the tacit sense of conscious unity does not have a phenomenal presence at all, that it does not make a phenomenal difference of any sort, then we are left to ask how this expectation could plausibly be identified with our sense of conscious unity. Our sense of conscious unity is a constant of experience during our wakeful hours. It is a real phenomenon—otherwise there would be no need to account for it—and so whatever is supposed to do the job of accounting for it must be sensitive to this fact. It is unclear how something that had no phenomenological impact whatsoever would suffice.

Another question for this identification of a certain expectation with the tacit sense of conscious unity has to do with the fact that many expectations seem to be momentary and short-lived, whereas the sense of conscious unity is pervasive and enduring. Here Rosenthal might distinguish between an expectation and its being present to consciousness. For instance, I expect that the oral defence of my dissertation will be a nerve-wracking experience, though the expectation is only occasionally at the forefront of my consciousness. This way of understanding expectations, however, will not help Rosenthal. There are many expectations we have that do not, when they are unconscious or in their “standing” state—one can think here of the difference between merely having the standing belief that $100 \div 5 = 20$ and having it consciously—that do not seem to have any experiential impact. For instance, I have the standing expectations that I can add 2 and 3 together in my head if I want to, that I can recall at will the births of my children, and that I can, if I wish, focus my attention on my big toe and wiggle it. But none of these expectations seem to have any bearing on my occurrent experience. Rosenthal could claim that the expectation in question is special in this regard, but at least one problem with claiming this is that it is not at all obvious that the expectation that I can become aware of

the bearer of my mental states is, in fact, special in this way. When I introspect on them, all of my standing expectations seem to be on a par with respect to what it is like, subjectively, to have them; it is unclear whether one is different from any of the others, including the standing expectation associated with the tacit sense of conscious unity. Of course, different expectations can have different behavioural and cognitive effects, but there does not seem to be much of a difference with respect to what it is like to have the different expectations in their standing states. This also makes it very difficult to say what the subjective difference would be between having the expectation (in its standing state) and not having it at all. I believe that these introspective observations justify the conclusion that expectations in their standing state do not have a phenomenal impact. And if it is correct that the expectation in question is like other standing expectations, then it is difficult to accept Rosenthal's identification of it with the tacit sense of conscious unity that pervades all of experience.

If the arguments I have presented in this section are sound, then the burden of accounting for our sense of conscious unity in a way that does not invite the Mismatch Objection falls squarely on the shoulders of the gathering strategy. Although the gathering strategy may escape the Mismatch Objection, it depends upon the possibility that a single HOT represent all of one's (conscious) lower-order states at once. Thus, it would seem, Rosenthal's entire account of conscious unity would seem to rest upon the representational power of HOTs.

CHAPTER 6

HUSSERL ON PRE-REFLECTIVE SELF-CONSCIOUSNESS

Up to this point, the focus has been on the phenomenal unity of consciousness. In this and the following chapter, the focus will shift to a different but related topic: self-consciousness. The topic of self-consciousness is a large one. My focus will be limited to consideration of a thesis I will call the “Ubiquity Thesis.” Just as theorists have debated whether consciousness is necessarily phenomenally unified, they have also debated whether self-consciousness is a necessary feature of consciousness; that is, they have argued that it is impossible for an individual to be in a conscious mental state and, at the same time, fail to be self-conscious.

Ubiquity Thesis Necessarily, if an individual is in a conscious mental state, that individual is also self-conscious.

Although there have been many theorists who have defended the Ubiquity Thesis, the focus in this chapter will be on the views of Edmund Husserl. Husserl’s work launched a philosophical tradition known as phenomenology. Many of the philosophers working within the phenomenological tradition maintain the Ubiquity Thesis (see, for instance, Heidegger (1982), Sartre (1956, 1957), and Henry (1973)) and so studying the views of Husserl is an important first step toward engaging with these philosophers. The Ubiquity Thesis has more recently been defended both by those identifying with the phenomenological tradition and by some who do not (see Smith, 1989; Flanagan, 1992; Kapitan, 1999; Zahavi, 2003b, 2005; Kriegel, 2004, 2009). I find the Ubiquity Thesis to be implausible, and in this paper I will argue against Husserl’s defence of

it. In the next chapter, I will argue against a recent defence of the thesis by Kriegel. Before taking up Husserl's views, it will be beneficial to make a few comments about the kind of self-consciousness invoked by the Ubiquity Thesis.

6.1 Self-Consciousness

Not all of the philosophers who have defended the Ubiquity Thesis have defended it for the same reasons or understood it in precisely the same way, but what is common to them all is that they have in mind a relatively thin or minimal notion of self-consciousness.

It is common to distinguish between more robust conceptions of self-consciousness and more minimal conceptions of the self-consciousness. Few philosophers or developmental psychologists make the distinctions in the same way, but there are a number of different properties or capacities that are commonly associated with a robust conception of self-consciousness. For instance, Butterworth, following Edelman (1989), maintains that "higher-order self-consciousness" occurs when an individual has a self-concept and when "the self is the object of one's own cognition" (1995, 36). Lewis (1992, 2003) suggests that "objective self-awareness" involves the capacity to recognize oneself in a mirror, to refer to oneself with first-person singular pronouns, and to experience certain self-evaluative emotions like pride and shame.¹ Baker (1998, 2000) argues that a "first-person perspective" requires the ability to think of oneself *as* oneself; that is, it requires the ability to conceptualize the distinction between oneself and everything else there is. Flanagan (1992, 195) says that "strong self-consciousness" involves the ability to think about "one's model of one's self." This can plausibly be understood as the capacity to think about one's personal identity, where "personal identity" is understood in its colloquial sense. To do this, one must be able to see

¹See also Duval and Wicklund (1972).

oneself as the main character of a narrative about a temporally extended subject of consciousness and agent of action in the world. These are all properties of a robust self-consciousness. None of the defenders of the Ubiquity Thesis claim that an individual need possess or be actively engaging any of *these* capacities in order to satisfy the Ubiquity Thesis.

“Self-consciousness” can also refer to a specific kind of mental activity: the activity of introspecting or reflecting upon one’s experience. For instance, when I am reading a book, my attention is directed “outward” and the focus of my attention is on what I am reading, but when I pause my reading and introspect or reflect on the experience of sitting in my chair and reading my book, my mind ignores stimuli from the external world and turns its attention “inward” to focus on its own states. This kind of self-consciousness is often referred to by phenomenologists and others as “reflective self-awareness.”² Again, the Ubiquity Thesis is not intended to invoke this kind of mental activity. Rather, the thesis concerns what phenomenologists call *pre*-reflective self-awareness.³

The above list of capacities and the characterization of the mental activity often thought of as introspection amount to a positive characterization of robust self-consciousness. To say that an individual is self-consciousness in this sense is to say that he or she has some of these mental capacities or that he or she is engaged in the mental activity of introspection. The minimal or thin notion of self-consciousness that defenders of the Ubiquity Thesis invoke is usually characterized negatively, i.e., as a kind of self-consciousness that does not involve any of the capacities or activities associated with a robust self-consciousness. Defenders of the Ubiquity Thesis offer comparatively little by way of a positive characterization of thin or minimal

²Kriegel (2004, 187) calls it “transitive self-consciousness.”

³Alvin Goldman (1970, 96) also uses this terminology. Kriegel calls it “*intransitive* self-consciousness. I find Kriegel’s terminology puzzling and take it up in section 7.2.1.

self-consciousness. One reason that is sometimes given for this is that it follows from the very nature of the sort of self-consciousness under discussion that it should be difficult to describe. They insist that minimal or thin self-consciousness has a phenomenological manifestation, but they are careful to point out that, compared to, say, phenomenally-given redness, it is phenomenologically unimpressive; it is subtle and runs in the background. Because of this, it resists positive description. This, in turn, can make it quite difficult to get a clear sense of what minimal or thin self-consciousness is. When I consider the views of Husserl and Kriegel below, I will present what they have to say about the kind of self-consciousness they believe to accompany all conscious states. For now, however, the negative characterization will do.

In what follows, it will often be convenient to drop the adjectives “minimal” and “thin.” Any unqualified occurrence of “self-consciousness” or related terms should be understood to refer to the sort of minimal or thin notion of self-consciousness just outlined. When I intend to refer to a more robust notion of self-consciousness I will be sure to indicate it.

6.2 Husserl and the Ubiquity Thesis

Husserl was a student of Franz Brentano. Brentano is, perhaps, most well-known for a claim about the intentionality of mental states. According to him, every mental state has, or is directed at, an object. The notion of an object, here, is similar to the grammatical notion of an object. Brentano’s claim was not that all of our mental states are directed at *physical* objects but, rather, that each mental state is directed at (or about) some object or other. Our mental states can be directed at physical objects, but they can also be directed at abstract objects or at mental objects like themselves. Brentano also argued that every mental state has *two* objects: a primary object and a secondary one. When I look at the mug on my desk, the primary object

of my mental state is the mug. But my mental state also has a secondary object: itself. Brentano called the way in which a mental state has itself as its secondary object its “inner consciousness.”

It is important to note that, although Brentano maintained that every mental state has itself as one of its objects, he did not maintain that the relation between a mental state and its primary object is exactly like the relation between it and its secondary object. The primary object of one’s mental state is the object one happens to be attending to at a particular moment. The secondary object of a mental state is not the object of attention. This difference between primary and secondary objects manifests itself in cognitive states. The primary object is, as phenomenologists say, “thematized” in cognitive states. Secondary objects, on the other hand, are not thematized in cognitive states; they do not appear as contents of those states. Although it is possible to direct one’s attention inwardly in an act of reflection or introspection, thereby making a (formerly) secondary object into the primary object, mental states are usually not the focus of attention.

Brentano’s claim that all mental states have themselves as objects can be seen as a particular way of accepting and unpacking the Ubiquity Thesis. Indeed, Kriegel’s (2009) view, which I will discuss in the following chapter, could be seen as a contemporary defence of Brentano’s view. What is relevant with respect to our discussion here, however, is that, although Husserl accepted and defended Brentano’s intentionality thesis, Husserl rejected Brentano’s claim that all mental states have themselves as secondary objects.

The natural response to this would be to think that Husserl therefore did not accept the Ubiquity Thesis. A mental state having itself as an object, albeit a secondary object, might count as a kind of self-consciousness. By rejecting this aspect of Brentano’s view, Husserl could be seen as rejecting one way of understanding the Ubiquity Thesis. Indeed, this is what the traditional understanding of Husserl has

been; most have taken him to maintain that the only sort of self-consciousness there is involves the deliberate turning of one's attention inward. Since this kind of self-consciousness happens only occasionally, we would not, as a matter of general course, be self-conscious. Zahavi has recently challenged this understanding of Husserl. Zahavi argues that, although Husserl rejected Brentano's views about the prevalence of a secondary object in all our conscious states, Husserl nevertheless affirmed the Ubiquity Thesis in a different way. It is this understanding of Husserl that I will examine in this chapter. I will not, however, concern myself with the question whether Zahavi's understanding of Husserl is the correct one. Instead, I will assume, for the sake of argument, that Zahavi's interpretation of Husserl's mature view is a fair one and ask the question whether Husserl's view presents us with reason to accept the Ubiquity Thesis.

6.2.1 Zahavi's Interpretation of Husserl

The traditional understanding of Husserl's view on self-consciousness is that Husserl understood consciousness strictly in terms of (primary) object consciousness and that, because of this, the only kind of self-consciousness Husserl countenanced was of the deliberate, reflective, and objectifying kind. In his earlier works, Husserl, in rejecting the idea that we are always reflectively aware of our conscious lives, often asserts that we are not aware of our conscious states in the way that we are aware of the objects we happen to be focusing our attention upon. In Husserl's terminology, we "live through" our experiences, the idea being that we are not, in general, aware of our experiences *as experiences*. Our attention is focused outwardly and so the fact that we are having experiences is not an explicit part of our conscious life. The inference to the traditional interpretation of Husserl is made relatively quickly. If the only kind of consciousness or awareness that Husserl accepted is modelled on our awareness of external objects and if he maintained that we are not aware of our own mental states

as a matter of general course, then Husserl seems compelled to reject the Ubiquity Thesis.

Zahavi (2003a, 2005) argues, however, that Husserl, at least in his later and posthumously published works, came to countenance an additional kind of consciousness or awareness. To support his claim, Zahavi marshals numerous examples from Husserl's manuscripts in which Husserl explicitly claims that we are aware of all of our experiences. In these texts, Husserl claims that, although we do not notice or pay attention to all of our experiences and we do not perceive them like we do external objects, we are conscious of all of them. One possible reaction to these claims is to take Husserl to be contradicting himself. Zahavi opts for a different interpretation. He attempts to read these assertions in a way that makes them consistent with Husserl's earlier claims about self-consciousness. This leads him to suggest that Husserl, at least later in his career, countenanced two kinds of awareness: a kind of awareness or consciousness that occurs when we focus our attention on an object (whether an external object or an internal mental state) and a kind of awareness that is not objectifying, reflective, or deliberate. The ubiquitous awareness we have of our conscious states is of the latter sort.

Aside from the texts in which Husserl explicitly asserts that we are conscious of all of our experiences, there is another reason to opt for Zahavi's interpretation of Husserl's mature view on self-consciousness: Husserl seems to argue that the non-objectifying, pre-reflective awareness we have of all of our conscious states is a necessary condition for *reflective* self-awareness. During an episode of reflective self-awareness, consciousness directs its intentional arrow at itself and takes itself as object. Taking something as object, is a mental *act*. And every mental act, according to Husserl, has a certain duality to it. In a mental act, the subject is both active and passive. The subject is active in that the subject "positions" itself with respect to the object. A subject can position itself with respect to an object (or objects) in a

number of different ways: by making a judgement about the object, by comparing or differentiating a pair of objects, by forming a wish or desire with respect to an object, or, in the most primitive sort of positioning a subject can do with respect to an object, by merely noticing or paying attention to it.⁴ But according to Husserl, being active presupposes a certain kind of passivity. It presupposes that there is something the subject is acting upon, that there is something that is given to the subject to respond to. Without something being given to consciousness, there would be nothing to affect the subject and provoke a response.

This duality that pervades all mental activity has implications for self-consciousness. In reflective self-consciousness, consciousness takes itself as its own object. This implies that the subject is being active, positioning itself with respect to itself. But as we have just seen, Husserl believed that consciousness positioning itself with respect to an object requires that the object be given or presented to consciousness. In the case of reflective self-consciousness, however, the object that is given to consciousness is itself. Thus, in order for consciousness to engage in an episode of reflective self-consciousness, consciousness would have to be already given to it. Herein lies the need for a non-reflective, non-positional self-consciousness. How could consciousness be given to itself to serve as its own object if it were not, prior to the act of reflection, already aware of itself in some way or other? Reflection does not *produce* its object. Rather, reflection is, in part, a *response* by consciousness to what is given to it beforehand. Thus, the phenomenon of reflective, objectifying, positional self-awareness presupposes another kind of self-awareness, a self-awareness that does not involve taking oneself as object and persists between and through episodes of reflective self-consciousness.

⁴This vocabulary, regarding a subject “positioning” itself with respect to its object, is taken up by many phenomenologists and manifests itself in discussions of self-consciousness when pre-reflective and reflective self-consciousness are alternately called “non-positional” and “positional” self-consciousness respectively.

Not only does Husserl argue for the existence and pervasiveness of pre-reflective self-consciousness, he also proposes an analysis of its structure. When we engage in reflection and introspection, the object of our attention, namely, our inner conscious life, is given to us as persisting through time and having temporal extension. One of the ways this manifests itself is that we automatically and intuitively see the experiences that we “live through” as being part of the same stream of consciousness as past experiences. This fact about how we intuitively regard our own experiences in reflection requires, according to Husserl, that they have a certain structure. Husserl’s account of this structure is known as his analysis of inner time-consciousness, and it piggybacks on his account of how we perceive objects that persist through time. Thus, I will take a few moments to recount his views about how our perception of these objects is structured.

6.2.2 Husserl’s Account of Inner Time Consciousness

Husserl begins with the observation that, if awareness at a time was restricted to an awareness of an object’s contemporaneous momentary stage, then our conscious lives would be much different than they are. It is doubtful, for instance, that we would ever perceive a present object of attention as being the same object that was perceived moments ago. As I attend to the coffee mug on my desk, I perceive it as the same object I looked at minute ago. Husserl maintains that this would not be possible if all I was conscious of, at the moment I look at my mug, was its contemporaneous momentary stage. I would (now) have no awareness of any of the previous temporal stages of the mug, because the mental states by which I was aware of those stages would themselves have faded into non-existence with the passage of time.

The point can, perhaps, be made even more strongly by considering our awareness of temporally extended objects like melodies and spoken or written sentences. Hearing a series of musical notes as a melody requires an awareness of the different notes along

with their rhythmic and pitch relations to each other. For instance, hearing an E by itself is subjectively different than hearing an E after a C. In the latter, the E is heard as a major third. Hearing an E by itself does not evoke the same subjective sense. In the same way, hearing a B as a leading tone requires hearing it within a context in which other notes have sounded, in particular, other notes that are part of the C-major scale. Hearing the B within a different context could give one the subjective sense that one was hearing the third degree of a G-major scale. Similar considerations apply with respect to spoken and written language. Understanding what has been communicated upon hearing or reading the last word in a sentence requires some awareness of what has come before it. All these things require that the objects of earlier states of awareness somehow remain present to the mind at later times.

Husserl accounted for this by arguing that our awareness has a certain temporal width to it. This is due to the fact that, at any given moment, our conscious awareness has a certain tripartite structure. Suppose I am watching a batter take pitches from a pitcher. Now consider my overall state of awareness when the batter is in mid-swing. My consciousness of the mid-swing location of the bat is, what Husserl called, a *primal impression* or *primal presentation*. A primal presentation is the part of my total conscious state whereby I am aware of the temporal stage of the scene that happens to be simultaneous with my present state of consciousness. But while I have a primal presentation of the bat in mid-swing, I also, at the same time, have a *retention* of the bat in its location immediately prior to its present location. Even though the retention is the state whereby I am aware of a previous temporal stage of the bat, the retention itself is co-present with my primal presentation of the bat in mid-swing. Thirdly, I am also conscious of where the bat will be in a moment from now, and I am conscious of the future stage of the bat via a *protention*. In the same way that my retention of the bat's earlier stage is co-present with my primal

presentation of the bat's current stage, my protention of the bat's future stage is itself co-present with both my retention and primal presentation. In this way, my overall current state of consciousness has a tripartite structure and a temporal width. Even though my current state of consciousness will survive but a moment and will fade out of existence with the arrival of the next moment in time, it contains within itself an awareness of what occurred in moments past and of what will occur in moments yet to come.

How does all of this apply to self-consciousness? As we have already seen, Husserl maintains that reflection is not productive. Reflection does not construct its object; it merely responds to it and discloses what is already there. Husserl also maintains that, when we turn our attention inward and reflect on our conscious states, they are given to us as all being part of the same conscious life or stream of consciousness. Since reflection does not change the mental states it takes for its objects, the elements that produce the sense that they (the mental states) all belong to the same stream of consciousness must themselves all be part of the same stream prior to reflection.⁵ This is where Husserl's account of inner time-consciousness is pressed into service. The tripartite structure of retention–primal presentation–protention (hereafter “RPP”) includes one further wrinkle that is aimed at satisfying this requirement.

According to Husserl, a retention of an object includes not only the object but also the primal presentation of that object. Suppose that, at t_1 , a subject has a primal presentation of an object, x . We might represent this state of affairs with the notation, $P(x)$. According to Husserl, at t_2 , when the subject's awareness of x changes from a primal presentation of x to a retention, the structure of the retention is not $R(x)$, but rather $R(P(x))$; what is retained in the retention is not just an awareness of x 's

⁵It should be pointed out that, according to Husserl, although reflection does not alter the states being reflected upon, reflection *does* modify one's stream of consciousness. There's a difference between having an experience of x and reflecting upon an experience of x . But reflection does not, as it were, *penetrate* the experience reflected upon and change its content or character.

t_1 stage, but an awareness of the mental state that was the primal presentation of x 's t_1 stage.

In the absolute passing-on, in the flowing process, the first primal impression becomes changed into a retention of itself, this retention becomes changed into a retention of this retention, and so on. But together with the first retention there is a new “now,” a new primal sensation, and the latter is combined continuously with the former in one moment in such a way that the second phase of the flow is primal sensation of the new now and retention of the earlier now; the third phase is again new primal sensation together with retention of the second primal sensation and retention of the retention of the first; and so on. (1991, 85–86)

We could chart this subject's mental states over time as indicated in Table 6.1. (The subscript in “ $P(x_1)$ ” represents the temporal stage of the object x . Thus, “ $P(x_1)$ ” represents the primal presentation of x 's t_1 stage. The chart omits protentions and assumes that t_1 marks the beginning of a period of consciousness for the subject.)

t_1	t_2	t_3
$P(x_1)$	$P(x_2)$	$P(x_3)$
	$R(P(x_1))$	$R(P(x_2))$
		$R(R(P(x_1)))$

Table 6.1. Consciousness Over Time

In this way, a certain kind of self-awareness is built in to the stream of consciousness. As consciousness flows from the present into the past, it always retains an awareness of its earlier states. This makes it the case that the elements for the sense of identity we get, when we reflect on our conscious lives, are present within the stream of consciousness even when we are not actively reflecting or introspecting. The structure of our inner temporal awareness also provides the structure of our pre-reflective self-awareness.

There are a pair of related questions that emerge from Husserl's account of time-consciousness and pre-reflective self-consciousness. First, Husserl's account of the inner flow together with the above notational representation of his account make it seem as though retentional states have only a single intentionality: namely, they intend previous mental states, not the objects of the previous mental states. Second, if pre-reflective self-consciousness is accounted for by the way in which retentional states retain primal presentations, then are our occurrent primal presentations included within our pre-reflective self-consciousness? In Table 6.1, for instance, the subject's primal presentation at t_2 is not intended by any retention at t_2 . Does that mean that, at t_2 , the subject's occurrent primal presentation is not included within his or her overall pre-reflective self-consciousness at the time?

These are good questions. The latter question, for instance, is used by Derrida (1973; 1990) to argue that, according to Husserl's account, we can only be aware of the present once it has passed. Husserl seems to have anticipated this worry. He asks, "What about the beginning-phase of an experience that is in the process of becoming constituted? Does it also come to be given only on the basis of retention, and would it be 'unconscious' if no retention were to follow it?" (1991, 123). His answer to this question is a resounding "no." "It is just nonsense," he says, "to talk about an 'unconscious' content that would only subsequently become conscious. Consciousness is necessarily *consciousness* in each of its phases" (original emphasis) (ibid.). This makes it clear that he sees what the potential problem is, but does he say anything that would make it clear why we should not think that his view has this problematic consequence?

Husserl claims that content is not made conscious by being taken in an act of apprehension. This would lead to an infinite regress, for the apprehending act would itself have to be conscious. Rather, "'content' is 'primally conscious' in itself and necessarily" (ibid.). The notational representation of his theory is not supposed to

be taken as having the consequence that primal presentations are not conscious or that we are only aware of them once they are past. What, then, is the *need* for retentions to be retentions of primal presentations rather than just of the objects of primal presentations? This has to do with facilitating the ability of consciousness to take itself as object. A primal presentation, argues Husserl, cannot be the (primary) object of consciousness as long as it is an occurrent primal presentation. It must first pass on into a retention. We can only focus our intentional gaze upon primal presentations that are preserved in retentions in the way that the theory prescribes.

What about the former of the two questions? Does the theory not have the consequence that we are aware of just-past primal presentations at the expense of being aware of the objects of those primal presentations? Husserl seems to suggest that consciousness can “decide” what it wishes to focus its intentional gaze upon. The intentional arrow of consciousness can rest either upon the retained primal presentation or “beyond” it upon what the primal presentation is a presentation of. The situation, he says, “is analogous to the re-presentation of an appearance of a physical thing, which possesses intentionality not only in relation to the thing-appearance but also in relation to the appearing thing; or better still, it is analogous to the memory of a memory of A, which makes us conscious not only of the memory but also of the A as what is remembered in the memory” (ibid., 86).

The traditional understanding of Husserl is one according to which he countenanced only one kind of self-consciousness: reflective self-consciousness. According to Zahavi, if we sample more broadly from Husserl’s later works and from unpublished manuscripts, we get a slightly different picture. We get a picture according to which Husserl understood conscious experience as consciousness being given to itself, a kind of pre-reflective, non-objectifying self-consciousness that persists through all of our conscious lives, even between times of active reflection and introspection. According to this view of Husserl, he argued for the existence of pre-reflective self-consciousness

by arguing that it is required by the possibility of reflective self-consciousness, and he provided an account of the structure of pre-reflective self-consciousness that dovetails with his account of inner time-consciousness. Thus, according to Zahavi's understanding of Husserl, Husserl affirmed the Ubiquity Thesis. In the rest of this chapter, I will raise some questions for Husserl's account.

6.3 Objections to Husserl

6.3.1 A General Worry

I want to begin my critique of Husserl's views on the pervasiveness of pre-reflective self-consciousness by considering an objection that was raised by a contemporary of Husserl's, H. J. Watt, and considered by Husserl in *Ideas* (§79). The objection is a sceptical one regarding our ability to learn about consciousness via introspection.

In the course of presenting his philosophical program, Husserl makes many claims about the nature of consciousness. Many of these claims are based on observations he makes while reflecting on his own conscious experience. Yet, as Watt points out, these claims are taken to reveal truths about the nature of conscious experience beyond reflection. Watt's worry is about whether the observations Husserl makes about his own experience while reflecting upon it can be taken to reveal anything about the nature of experience that is not reflected upon. The only access we have to experience outside of reflected experience is via reflection. We can never examine unreflected experience by itself, for when we attempt to do so, we engage in the activity of reflection. When not engaged in reflection, we "live through" our experience; our attention is focused upon whatever the primary object of our current mental state happens to be. We can only observe what experience is like when we deliberately turn our attention inward. Whatever features of experience reveal themselves to us in this act of inner attention are features of the experience as reflected upon. Whether

they are also features of experience when our intentional focus is directed *through* experience to external objects is not something we can assume.

Husserl does not say whether Watt's objection is motivated by a particular claim or set of claims that Husserl makes about unreflected experience or whether Watt is merely raising a general sceptical worry about Husserl's project as a whole. Whatever the case, it is clear that this objection has relevance to our discussion here. Husserl maintains that all experiences, even those that populate the significant periods of time in between moments of introspection exhibit a kind of self-consciousness. This claim is justified, in large part, by observations that are made during reflection. Indeed, Husserl's entire philosophical project proceeds on the basis of reflections upon experience. Thus, if the sceptical worry is a sound one, the implications for Husserl's claim about the ubiquity of self-consciousness are clear.

Husserl dismisses this objection. He argues that those who press the objection are themselves guilty of making assertions that, by their own rights, they have no business making. They themselves frequently make assertions about the nature of pre-reflected mental states. But if their scepticism is justified, then their own claims are undermined. Husserl says,

Every genuine scepticism, whatever its type and orientation may be, can be recognized by this fundamental absurdity, that in the arguments it uses it presupposes implicitly, as the conditions of the possibility of its validity, precisely that which it denies in its own theses. . . . He who merely says, I doubt the significance of reflexion for knowledge, maintains an absurdity. For as he asserts his doubt, he reflects, and to set this assertion forth as valid presupposes that reflexion *has* really and without a doubt (for the case in hand) the very cognitive value upon which doubt has been cast, that it does *not* alter the objective relation, that the unreflective

experience does *not* forfeit its essence through the transition into reflexion (original emphases). (1958, 227–8)

There are several things to say about Husserl's response to Watt's objection. Taken as a *tu quoque*, Husserl's response deserves to be taken seriously. One cannot question the merit of Husserl's claims about unreflected experience while using introspection to make certain generalizations about experience oneself. If introspected experience has limited value for Husserl, then it also has limited value for everyone else who speculates about the general nature of experience. Whether Watt, Husserl's direct target, was guilty of making such a self-undermining objection is not knowable from what Husserl himself says about Watt's views. But regardless of what Watt may or may not have been guilty of, Husserl's response to this objection should give anyone who wishes to raise this kind of sceptical worry pause.

This point is well-taken. However, there is a real sense in which Husserl's response falls flat. Pointing out that the sceptical worry applies as much to others as to himself does little to diffuse the challenge that it poses to his own project. If reflected experience is unsuited to justify claims about unreflected experience, or even if we are just not in a position to *know* whether our observations about reflected experience provide a firm footing for venturing claims about unreflected experience, Husserl's views on unreflected experience and experience in general become much more speculative in nature.

Husserl acknowledges that “*every variety of ‘reflexion’ has the character of a modification of consciousness*” (original emphases) (1958, 219). However, in Husserl's view, this “modification” does not adversely affect our ability to learn things about experience. Reflection modifies experience in that the stream of consciousness has a different character during a moment of introspection than it would have had if the subject had not been introspecting at the time. For instance, the character of my stream of consciousness while looking at my mug is different than it would have been

if I had introspected my experience while looking at the mug. Reflection alters one's overall phenomenal experience at a time, but, in Husserl's view, it does not alter the state being reflected upon. Although reflection has access to the experience its intentional gaze targets and has the ability to examine it, the intentional gaze does not alter the content of its target. The state that is introspected is not itself changed by being observed. Thus, although introspection adds to, and thereby modifies, one's overall experience at the time of introspection, it does not modify its target state. Given the nature of the objection being considered, however, Husserl's view about the way in which reflection leaves its target states unchanged will seem like nothing more than a naïve confidence unless it is justified in some way. Although Husserl has things to say about this, he does not offer a general argument for why reflected experience can reliably serve as ground for assertions about unreflected experience.

To some extent, Husserl can be excused for this. Nowadays, it is quite common for psychologists, neuroscientists, and philosophers alike to treat with suspicion the verdicts rendered by introspection and reflection. We have access to a century of empirical research on the mind and brain that Husserl did not have access to. We know that there are numerous examples of ways in which our subjective sense of our own experience is often seemingly inconsistent with what other methods of enquiry suggest about our mental and brain states. We also have much less riding on the veracity of the results of introspection than Husserl did. Husserl's grand philosophical project depends on the veracity of introspection. He argued that the proper domain of philosophy was the study of experience and that the only way to properly study experience was to engage in different sorts of reflection. The way in which theorists today attempt to integrate and account for the findings of empirical research within their considered views of the mind is something that Husserl would have resisted. Philosophy, in his view, is first and foremost a science of consciousness and the proper way to study consciousness is not by gathering large sets of data in laboratories but by

suspending one's beliefs about the relation between experience and the external world and focusing strictly upon what the experience is like from the inside (a technique Husserl called *epoche*). Within this conception of the subject matter of philosophy and its methodology, doubts about how the results of introspection generalize can have no place. Although Husserl can be given some slack for being as confident as he was in the deliverances of introspection, this sceptical worry is a real one and not one that Husserl has given an ultimately satisfactory response to.

There is another way in which Husserl's response could be seen to be problematic. In the quotation above, Husserl seems to imply that cognitive processes, such as reasoning and the drawing of inferences, are kinds of reflection and introspection. Speaking about his opponent, he says, "as he asserts his doubt, he reflects, and to set this assertion forth as valid presupposes that reflexion *has* really and without a doubt (for the case in hand) the very cognitive value upon which doubt has been cast..." (ibid.). The worry raised by his opponent is that we lack good reason to believe that the attributes we discern in reflected experience are also automatically present in unreflected experience. But asserting this kind of worry is not just a matter of intuiting something on the basis of one's experience. Husserl's objector is not claiming that a difference between reflected and unreflected experience is itself experienced or intuited. Rather, the objection involves engaging one's powers of reason and the ability to evaluate argumentation. This is what Husserl seems to characterize as reflection.

Husserl is free to define "reflection" in whatever way he pleases, and, indeed, Husserl catalogues a host of different kinds of mental activity as different kinds of reflection, but it seems open to Watt and to Husserl's other targets to respond to him by restricting the scope of their sceptical worry. One might well admit that the epistemic reliability of reason and inference can be questioned, but one need not raise these kinds of questions to press the sceptical worry against Husserl. One might

well suppose, for the sake of argument, that our powers of reason and inference are in relatively good epistemic standing and still question whether what we observe or sense about our own experience when we reflect on it can be assumed to carry over to experience that is not reflected upon. Calling into question our powers of reason invokes a rather extreme and far-reaching scepticism. If this kind of scepticism were pressed against Husserl, his response would seem to be spot on. But Husserl's opponents need not have anything quite so radical in mind.

The epistemic value of introspection is a large subject and there will be no attempt in the following pages to render a general verdict about the value of the data gleaned by introspection and reflection. Of interest, however, will be the arguments Husserl makes in justifying the existence and pervasiveness of pre-reflective self-consciousness. Some of these rely crucially upon introspection. I will argue that Husserl's views on pre-reflective self-consciousness should not be accepted. My arguments will, on occasion, proceed from what I take to be true about my own experience and will, in some cases, either implicitly or explicitly, be offered as justifications for claims about what unreflected experience is like. I offer these arguments with an eye toward the sceptical worry raised by Watt and the *tu quoque* response by Husserl. But whereas Husserl had little patience for Watt's objection and took the lesson from the *tu quoque* to be that we could be confident in whatever conclusions we drew from our first-personal investigations, I believe that the correct response is to regard all of our conclusions on these matters with less rather than more confidence. I believe that there is no such thing as pre-reflective self-consciousness, but, unlike Husserl, I believe that conclusions on this matter must be tentative.

6.3.2 Double Intentionality

Another objection that has been raised against Husserl's account is that it may be subject to the very same worries that he himself raised against Brentano. There

are several places where Husserl talks about the double intentionality of retentive states. Husserl even goes so far as to use similar language as Brentano; for instance, Husserl says that mental states not only have primary objects but also *secondary* objects. As Zahavi (2005, 60) rightly points out, these passages seem to indicate that Husserl forgot, at least momentarily, the problems he saw in Brentano's double-object account. Husserl objected to Brentano's view, because he thought it was tantamount to claiming that we always have a perception-like awareness of our own mental states. Indeed, this is what led him to claim, as we saw earlier, that it is only in reflection and introspection that we take our own mental states as objects. By claiming, as he later did, that mental states have themselves as *secondary* objects, Husserl seems to open himself up to the same worries he raised for Brentano.

Although Husserl's comments about the secondary objects of mental states certainly adds a wrinkle to producing an internally consistent interpretation of his views, Husserl's defenders have argued that there are enough resources in Husserl's writings to dismiss this worry. Zahavi, for instance, has argued that Husserl's account of the RPP structure of inner time consciousness and pre-reflective self-consciousness can be understood in two ways: one he calls the "Internal Object Account" and the other he calls the "Urbewußtsein Account" ("Ur" meaning "Original" and "bewußtsein" meaning "consciousness").

The Internal Object Account, as its name suggests, is one according to which self-consciousness, even pre-reflective self-consciousness, involves taking one's own mental states as objects. Zahavi (2005, 59–65) admits that there is some textual support for this understanding, particularly in Husserl's *Bernau Manuscripts*, but he proceeds to argue that this interpretation of Husserl should ultimately be discarded. Although there are places where Husserl flirts with what appears to be a Brentanian account of pre-reflective self-consciousness, Husserl also frequently says things that explicitly distance himself from such an interpretation. For instance, when Husserl talks about

the secondary objects of mental states, the term “secondary” is frequently placed within scare quotations. Husserl also argues that, contrary to what is the case with Brentano’s view, the kind of secondary object-taking he (Husserl) invokes does not generate a vicious infinite regress. Finally, even though the notation used to represent Husserl’s RPP account suggests that retentions take primal presentations as (primary) objects, this would be to misunderstand the account. For instance, in contrasting retentions and memories (or, as Husserl called them, “secondary memories” or “recollections”), Husserl is explicit in claiming that retentions are not objectifying. Memory, he says, “constitutes... an immanent or transcendent enduring objectivity,” but retention “produces no enduring objectivities (either originally or reproductively) but only holds in consciousness what has been produced and stamps on it the character of ‘just past’” (1991, 38). Retention may well preserve an intentionality or the taking of an object that was already part of a primal presentation, but it does not add an object.

The Urbewußtsein Account is one that attempts to understand Husserl’s views on pre-reflective self-consciousness in a way that does not involve consciousness taking itself as object. It does so by understanding pre-reflective self-consciousness as the *self-givenness* of consciousness. Pre-reflective self-consciousness is not a matter of consciousness turning its intentional gaze upon itself. Rather, pre-reflective self-consciousness is consciousness presenting itself to itself. Before evaluating this way of interpreting Husserl’s claims regarding the double-intentionality of retentive states, it will be necessary to say a few things about what is meant by the self-givenness of consciousness.

The notion of an object being given to consciousness is a central notion within Husserl’s views on perception and consciousness. Being given to consciousness is a necessary condition on a thing becoming the object of an intentional state. When it comes to ordinary objects in the physical world, they are given to consciousness via

the sensory content of intentional states. Although sensory content is not enough, by itself, to endow a mental state with intentionality, sensory content constrains the meaning (or “Sinn”) that the state can have.⁶ Thus, while the intentional gaze of the mind is directed upon things in the external world, it is the sensory content of perceptual states that present objects to consciousness and make them available to be taken as objects. Because being given to consciousness is a perfectly general requirement for becoming the object of an intention, consciousness must itself be given to consciousness if consciousness is to be eligible to become the object of an intention. Reflective self-consciousness is a matter of consciousness taking itself for an object and, since reflective self-consciousness is something that we as conscious creatures often engage in, consciousness must, as a matter of general course, be given to itself.

The way in which consciousness is given to itself, however, differs in significant ways from the way in which physical objects are given to consciousness. Physical objects “appear” to consciousness. That is, they are given in a *perspectival* manner. No physical object is ever given to consciousness completely all at once. This is true for many reasons but can, perhaps, be seen most simply by considering the fact that whenever one looks at a physical object, one only ever sees the half of the object that is on its nearer side. One can, of course, change one’s position with respect to the object, but then, when the previously unseen portion of the object is seen, the previously seen part of the object becomes unseen. Consciousness, on the other hand, is given to consciousness *absolutely*. It is, according to Husserl, senseless to talk about consciousness being given to itself in a perspectival manner; there is no “side” to consciousness or “part” of consciousness that remains hidden when it is presented. This is not to say that we are *infallible* when it comes to knowing our mental states.

⁶Husserl leaves open the possibility of a creature possessing states with sensory content but not interpreting any of it.

To say that consciousness is given absolutely is just to say that what is presented, when it is presented to itself, has no part that remains inaccessible.⁷

There is another way in which there is a stark contrast between the way in which things in the external world are given and the way in which consciousness or experience is given. External objects are given via sensory content and via sensory experience, but experience itself is not given via experience. External objects are given to consciousness via mental states that have sensory content and these mental states are distinct from the things they present to consciousness, but an experiential state is not given to consciousness via a distinct sensory state. This would be tantamount to claiming that we are aware of our internal states in the same way we are aware of external objects. The problems with this kind of view are numerous and have been well-documented.

The question then becomes whether being taken as an object of an intention is a necessary condition upon being given (to consciousness). We have seen that being given to consciousness is a necessary condition for becoming the object of an intention. The question is whether the requirement goes both ways. With respect to things in the external world, the answer seems to be no. Husserl talks about the fact that ordinary perception usually involves a “background.” He considers a case in which he looks at a sheet of paper.

Around and about the paper lie books, pencils, ink-well, and so forth, and these in a certain sense are also “perceived,” perceptually there, in the “field of intuition”; but whilst I was turned towards the paper there was no turning in their direction, nor any apprehending of them, not even in a secondary sense. They appeared and yet were not singled out, were

⁷Just as Husserl’s confidence in introspection seems somewhat misplaced nowadays, this claim about consciousness being given absolutely is likely to seem somewhat naïve. I will later suggest that, outside of reflection, consciousness may not be given to itself at all, and so I will leave unchallenged the point about being given absolutely.

not posited on their own account. Every perception of a thing has such a zone of *background intuitions* (original emphasis). (1958, 117)

Husserl goes on to say that, although this situation involves a kind of awareness of the things that comprise the background, the awareness is not of the full-blown, objectifying variety. According to the Urbewußtsein Account, the same is true of consciousness itself. While engaged in ordinary acts of external perception, consciousness remains given to itself, available to be taken as the object, but, until such time as an act of reflection occurs, is not taken as an object.

In the rest of this chapter I want to take up, in turn, Husserl's claim that consciousness is always given to itself and the account that Husserl gives of it, i.e., his account of inner time consciousness.

6.3.2.1 The Self-Givenness of Consciousness

If we interpret Husserl charitably, Husserl rejects the idea that the kind of double intentionality inherent in retentional awareness involves consciousness taking itself as object. Pre-reflective self-consciousness is to be understood, instead, as the self-givenness of consciousness. In this section, I want to question this way of conceiving of pre-reflective self-consciousness.

The trouble for this way of conceiving of pre-reflective self-consciousness arises, in my view, from a consideration of the relationship between the givenness of consciousness, reflective self-consciousness, and what Husserl calls the “natural attitude” or “natural standpoint.” During the normal course of our day-to-day lives our intentional focus is on the things in our environment. Our minds are directed toward the physical objects we interact with and toward the mundane tasks and activities that are a part of our lives. The natural standpoint is our default mental attitude or posture. It is, however, possible for us to step away from the natural standpoint into another. When I turn my mind toward contemplating mathematics and arithmetic,

for instance, I adopt the arithmetical standpoint (Husserl, 1958, 104). While in the arithmetical standpoint, a whole new world of objects, consisting of things such as numbers and the law-like relations that hold between them, is available to me. The natural world remains available to me from within the arithmetical standpoint. However, “*the arithmetical world is there for me only when and so long as I occupy the arithmetical standpoint* (original emphasis)” (ibid.).

There are other standpoints that one can step into from the natural standpoint. Indeed, some of them feature crucially in his description of the methodology that is part of Husserl’s grand phenomenological project. In order to help his readers understand the proper subject matter and starting point from which his enquiry begins, Husserl describes several “reductions.” These are not ontological reductions. Instead, they are steps that successively restrict one’s focus until one has reached the point of contemplating pure consciousness. The first reduction that is performed in this process is what some have called the “psychological reduction” (Smith and McIntyre, 1982, 95). This reduction is, in essence, the most basic kind of reflection that one can do upon one’s own mental states. One merely turns one’s attention away from things in the external world to one’s own mind. This is nothing more than engaging in simple reflective self-consciousness, and doing so involves adopting a mental stance or attitude that differs from the natural attitude.

The question here is whether the kind of attitude that is taken up in the psychological reduction is to be understood as being similar to the arithmetical attitude in certain important respects. For instance, of the arithmetical attitude or stance, Husserl seems to claim that the objects that populate the realm of mathematics are only available to the mind from *within* the arithmetical standpoint. Thus it would seem that, in order for one to have an intention with a mathematical entity for its object, one would have to be in the arithmetical standpoint. But in addition, it would seem that mathematical objects are only *given* to the mind from within the

arithmetical standpoint. Although it is easy enough for the mind to move from the natural to the arithmetical standpoint, and, in this way, the mathematical objects could, in a sense, be said to be available to the mind even in the natural standpoint, the objects themselves are only given, in the proper sense, to the mind from within the arithmetical standpoint.

Are we to understand the psychological reduction in the same way? Certainly the mind can freely move back and forth between the natural standpoint and the standpoint of psychological reflection, and, as a result, one might say that the potential objects of intentional focus from within the reflective standpoint are, in a sense, available to the mind from within the natural standpoint, but are they *given* to the mind, in the proper sense, within the natural standpoint? The answer to this question is important with respect to the Ubiquity Thesis. For although we can freely (and often do) enter the reflective stance, if our own mental states are not given to consciousness outside of the reflective stance, then the Ubiquity Thesis would not hold.

The defender of Husserl is likely to point out that, when it comes to the givenness of mental entities, there is some textual evidence that would suggest Husserl did not think of the arithmetical standpoint and the so-called psychological reduction as being analogous in the relevant way. For instance, in *Ideas*, Husserl says that the Ego, consciousness, and experience are all “given to us from the natural standpoint” (1958, 112). Although quotations like this cannot be ignored, I also believe that a quotation like this is not, by itself, enough to put the worry I have raised to bed.

There are two reasons for this. First, the primary focus for Husserl, in the context of this quotation, is not upon cataloguing which objects are given from within different standpoints. Instead, Husserl is embarking on a description of what is available for consideration after all the relevant reductions have been made. He is concerned primarily with the “phenomenological residuum” that is left after all of the reductions, including the final “transcendental epoche,” have been made. By the time the

transcendental epoche has been reached, the natural standpoint has long since been abandoned. I believe it is permissible to read the relevant quotation above as a kind of off-hand comment.

Second, even if Husserl was self-consciously asserting that consciousness is given within the natural standpoint, there would still remain the need for Husserl to explain how this position should be reconciled with his comments about the givenness (or lack thereof) of mathematical objects in the natural standpoint. Both mathematical contemplation and reflection upon one's own mental states require a turning of the attention away from the external world. Both seem to require the acquisition of certain concepts and mental capabilities. I cannot, for instance, contemplate numbers or their relations to each other unless I have the concept of number, and I cannot reflect on my own mental states and entertain thoughts about them unless I have the concept of a mental state. Thus, both the arithmetical and reflective standpoint require the mind to engage some relatively sophisticated capabilities. Although more mature human cognizers are able to step back and forth between these standpoints with relative ease, the fact remains that the divide the mind must step over each time the transition is made to one of these standpoints is a significant one and one that could certainly, on the face of it, provide the basis for a difference with respect to which objects are given to consciousness from the natural standpoint and from within the arithmetical and reflective standpoints respectively.

Husserl and his defenders might respond by arguing that there is a clear difference in the way that mathematical objects and mental states are given to consciousness. Mathematical objects are ethereal, abstract entities. They are never present to consciousness in the vivid and direct way that, for instance, sensory mental states are. When a mathematical object is not actually the object of an intentional state, it seems to be completely absent from consciousness. However, even when we are not engaging in active reflection, we still enjoy conscious mental states. As long as we are

not asleep or otherwise unconscious, we are always in the presence of conscious mental states. This, the defender of the Urbewußtsein Account might say, is exactly what is meant by the self-givenness (i.e., the given-to-self-ness) of consciousness. Whenever there is something it is like *for* the subject to be in a particular mental state, then that state is given to consciousness. Thus, although there are similarities between the arithmetical and reflective standpoint in that they both require certain relatively advanced cognitive and mental machinery in order to apprehend the entities in their respective domains and make them objects of intentions, there is also a significant and relevant dissimilarity that could be used to explain why experiences and mental states are given within the natural standpoint but mathematical entities are not. There is more to an object being given to consciousness than a mere possibility that it might become an object of consciousness. When I am in the natural standpoint, both the number 2 and my occurrent conscious state of my computer monitor enjoy the possibility of becoming objects of an intention of my. However, only my consciousness of my monitor is “there for me” while I am still in the natural standpoint.

It is undeniable that we have the sense that our experiences are constantly there for us during our wakeful hours. It is also undeniable that we have the sense that we can turn our attention inward and reflect on our conscious states at any point. For these reasons, the claim that our experiences are always given to consciousness, even when we are not actively reflecting upon them, may seem unassailable. This claim, however, is embedded in a way of thinking about what is given to us in experience that does not have to be accepted. There is a perfectly coherent account of what is given to consciousness, both from within and from without the natural standpoint, that can explain why we have the sense that our experiences are always given to us and according to which our experiences are not self-given as a general matter of course.

External physical objects, arithmetical entities, and mental states are all alike in that, in order for them to become (primary) intentional objects, consciousness must enter a certain standpoint. While within the natural standpoint, our attention is fixed on the things in the world around us, and it is these things that become the objects of intentions. It is only when we step out of the natural standpoint and into another that objects we do not perceive with our senses can be taken as intentional objects. In order to contemplate mathematical entities, for instance, our minds must partially disengage from things in the external physical world and engage with a different set of objects, objects that are not perceivable or made available to consciousness via sensation. The same is true about our own mental states. In order to engage my own mental states *as mental states*, that is, in order to reflect on them, think about them, or form beliefs about them, I must disengage from the external world somewhat so that my mind can adopt a certain stance with respect to them.

What physical objects, on the one hand, and arithmetical entities and mental states, on the other hand, do not have in common is their respective givenness to consciousness from outside of the natural standpoint. As long as I am in a wakeful state and my sensory systems are functioning properly, the objects I perceive via sensation are given to consciousness. It is unobjectionable that physical objects are given to me while within natural standpoint. The only question about the givenness of these objects while within the natural standpoint concerns the objects in the perceptual background. I am inclined to agree with Husserl and say that they are given to consciousness even if they are not the occurrent objects of intentions. What about the givenness of external objects from outside of the natural standpoint? Here again, I am inclined to agree with Husserl and say that the objects we perceive via our sensory organs are given to consciousness even when consciousness has stepped out of the natural standpoint. We do not stop having sensory experience when, for instance, we begin doing arithmetic. It may be the case that, while we contemplate

mathematical objects, none of the objects in our sensory fields are taken as the objects of intentions, but the perceptual background persists and so there is good reason to maintain that the objects in our sensory fields are given to us from outside of the natural standpoint.

Arithmetical entities and mental states, however, are not given to consciousness outside of the mental standpoints that one must adopt in order to take them as intentional objects. As was pointed out above, this claim seems more plausible with arithmetical entities than it does with mental states: mental states seem to be “there for me” in a way that arithmetical entities do not seem to be. First, why believe that our own mental states are *like* arithmetical states in that they are not given to consciousness outside of the natural standpoint? The reason for this has been hinted at above. The way in which arithmetical entities and our own mental states become the objects of intentions are very different than the way external physical objects become the objects of intentions: in both cases, the mind must disengage from the sensory realm to some extent and engage a different set of mental processes and abilities to do so. Second, why is it so easy to believe that our own mental states are always given to consciousness, even when we are not actively reflecting or introspecting? This is because it is easy to equate the givenness of objects in the perceptual background with the givenness of mental states. It is easy say that mental states are always “there for me” because it is easy to mistakenly take the constant givenness of the objects in the perceptual background to be the givenness of my own mental states. What is always “there for me,” while in my wakeful hours, are the objects in my sensory fields, not my own mental states.

Husserl frequently speaks about the way in which we “live through” our experiences while in the natural standpoint. This expression carries with it the connotation of transparency, and it evokes certain metaphors. Windows are transparent. When we look through them at what is beyond, it is as if they are not even there. It takes

effort to fix one's gaze at the glass itself. One can, of course, turn one's attention to the glass at any time, but the glass is not given to the perceiver in the way that the objects beyond the glass are. The same is true of our experiences. When we "live through" them, it is as though they are not there. Provided we have the required mental abilities, we can at any time turn our attention to our mental states, but they are not, in general, given to us in the way that the object we perceive via them are.

6.3.2.2 Retaining Consciousness

In this last section I want to raise some questions for Husserl's account of inner time-consciousness. Husserl's account, you will recall, endows retentional states with a double intentionality. Primal presentations are the means by which we are aware of the now-stage of the thing our attention is focused on. When primal presentations flow into the past, they are modified into retentions. Retentions have a double intentionality in that they retain not only the temporal stage of the object of attention but also the primal presentation whereby that stage of the object was presented to consciousness. As we saw above, it has been argued that this double intentionality is not to be understood in a way that commits Husserl to the view that consciousness takes itself *as object* as a general matter of course. The questions I want to raise for Husserl's view of inner time-consciousness have to do with the reasons Husserl gives for thinking that our consciousness must be structured in this way. One of Husserl's motivations for this particular view of inner time-consciousness is his belief that, without the posited double intentionality, our retentional states would not provide us with the sense we have, when we reflect and introspect, that our reflected states are part of a temporally extended stream of consciousness; we would not have the sense that our reflected mental states are all part of the same stream of consciousness, even when those reflected states are separated in time. It is this necessary connection be-

tween our sense of the temporal extension of our experiential stream and the double intentionality of our retentional states that I want to question.

It is, perhaps, easy to see how we would come to have the aforementioned sense of our experiential stream if the temporal flow of our mental states had the structure Husserl claims for them. Consider again the subject's overall conscious state at t_3 in Table 6.1. It includes both a retention of the primal presentation of x_2 and a retention of the retention of the primal presentation of x_1 . Because of this, the subject's overall state of consciousness at t_3 includes traces of its consciousness from earlier times. Husserl likens this feature of our inner time-consciousness to that of the tail of a comet. More importantly, the remnants of earlier primal presentations appear in a format that encodes their temporal relations to each other; because the primal presentation of x_2 is retained alongside a retention of *the retention of* the primal presentation of x_1 , the primal presentation of x_2 can be readily interpreted to have occurred after the primal presentation of x_1 . And because the temporal order of the retained states can be "read off" of their structure, it is also relatively easy to see how it could make possible our awareness of temporally extended objects like musical melodies and events like the swinging of a bat.

The mere fact, however, that a certain structure, if it were instantiated, could account for a pair of phenomena that we desire an account of does not show that the said structure is *necessary* to account for those phenomena. What I will do now is offer another account of the structure of our inner time-consciousness and argue that a temporal progression of mental states that instantiated this structure could also account for the phenomena that we need an account of. The account I will offer, however, does not involve the kind of double intentionality that Husserl's account involves.

On the account I am offering, the only thing retained by retentional states are the temporal stages of the primary objects of the subject's awareness. Thus, if I am

aware of my mug at t_1 , then, at t_2 , the retentional state I occupy merely retains the mug's t_1 stage; it does not also retain the primal presentation of the mug's t_1 stage.⁸ On my view, a chart of the temporal progression of a subject's states of awareness would look like the one in Table 6.2. On Husserl's account, retentions are complex and

t_1	t_2	t_3	t_4
$P(x_1)$	$P(x_2)$	$P(x_3)$	$P(x_4)$
	$R(x_1)$	$R(x_2)$	$R(x_3)$
		$R(x_1)$	$R(x_2)$
			$R(x_2)$

Table 6.2. Consciousness Without Double Intentionality

contain within them an indefinite continuum of embedded retentions. On my account, retentions do not, as a general matter of course, embed mental states.⁹ Rather, they retain the temporal stages of the objects of earlier states of awareness. As retentions age, they slowly fade from consciousness.

Can this model account for our awareness of temporally extended objects? I believe it can. On Husserl's account, the temporal ordering of mental states is represented via the relative depth of their embedding in retentional states. On my account, the temporal ordering of primal presentations and retentions is represented in the structure via the vividness of the retention (as indicated by the reduction in size of the lettering as the retention of a particular temporal stage of the object of attention fades slowly from consciousness). This does not mean that, on my account, retentions are to be thought of as dimmer or less-vivid primal presentations or as a primal presentation with its brightness turned down a notch. If retentions were just dimmer primal presentations, then, for instance, our visual experience of moving

⁸There is an exception. If the subject happens to engage in reflection or introspection, then the primary object of attention will be a state of consciousness. At such a time, the retention would then be a retention appropriate to the object of attention.

⁹I say "as a general matter of course" because retentions will embed mental states during periods of reflection and introspection.

objects would be similar to timed exposure photographs; a visual experience of a batter swinging a bat would literally be a blur of the bat in the different stages of its swing across the visual field. Retentions preserve for consciousness an awareness of past stages of objects, but not by including those temporal stages into the overall phenomenal content of my occurrent experience. My occurrent experience consists only of the totality of what I am aware of via my primal presentation. Retentions give rise to my occurrent sense of motion, but I do not literally see the motion in my occurrent experience. As time passes and retentions slowly fade from existence, they remove their respective objects from consciousness as well. This corresponds, in the case of visual experience of motion, to the fact that our sense of motion also has a temporal limit into the past.¹⁰ Of course, we can often go on to recall a particular event or temporal stage of an event in memory or recollection, but this kind of mental event, though it occurs within the stream of consciousness and involves primal presentations and retentions, is not merely the reviving of an extinguished retention.

What about our sense, when we reflect, that the states we reflect upon are part of a temporally extended stream of consciousness? I believe my model can account for this just as well as Husserl's. On my account, as in Husserl's account, our primal presentations are accompanied by indefinitely many retentional states. Having the previous stages of the object retained in awareness is what gives me the sense, now, that the object I am now aware of is the same object I was aware of moments ago. This is what gives me the sense that the object preserves its identity over time. I believe that the sense that one's reflected state is part of a temporally extended stream can be given a similar explanation. Say I turn my attention inward to reflect upon my

¹⁰I should say, here, that in speaking about the object of a retention, I am promoting a certain way of thinking about retentions that is at odds with Husserl's view. According to Husserl, retentions are not intentional states *by themselves*. Rather, they are components of complex states of which certain other parts are responsible for providing the intentional content. I am not here intending to signal a difference between my view and Husserl's. I am merely speaking about the object of retention for sake of ease of communication.

experience. In doing so, my intentional gaze will land upon a mental state from a particular point in time. But at any given time, one's overall mental state consists of a primal presentation and an indefinite number of retentions. When my intentional gaze lands upon a mental state from a particular point in time, it will be presented with a package of retentional states: namely, all of the retentional states I was in at the time in question. The sense that the overall state I reflect upon is one that is a part of temporally extended stream is generated by the fact that, when I reflect, I become aware of the numerous retentional states I was in at the time in question. Husserl embedded primal presentations within retentions, because he believed that that was the only way he could satisfy the requirement that our mental states must be given to us in order to be taken as objects. But as long as one's account makes mental states available to be taken as intentional objects when consciousness turns its attention to them, Husserl's requirement can be satisfied. When we introspect, our minds adopt a certain stance with respect to our own mental states. There is no reason to believe that, once this stance is adopted, consciousness cannot take retentions, rather than merely their contents, as objects of intentional states.

Thus, it would appear that my model is able to account for the same phenomena that Husserl's account was constructed to account for. What is important for the argument I am making here, however, is the fact that, if my model is successfully able to account for the relevant phenomena, then Husserl's claim that inner time-consciousness *requires* a certain structure would be shown to be false. The success of my model would also have further import. The inherent double intentionality of Husserl's account of inner time-consciousness is a natural fit for the claim that pre-reflective self-consciousness is ubiquitous. If my model is able to do the explanatory work that Husserl's is able to do, then my model undermines the Ubiquity Thesis in an indirect way by removing one of the planks that one might use to build a consensus case for the Ubiquity Thesis.

CHAPTER 7

KRIEGEL ON INTRANSITIVE SELF-CONSCIOUSNESS

In the previous chapter, I presented Zahavi's interpretation of Husserl and argued that Husserl does not give us good reason to accept the Ubiquity Thesis. In this chapter, I will examine a contemporary theory of consciousness that entails the Ubiquity Thesis. The theory in question is that of Kriegel.

There are several reasons why a consideration of Kriegel's theory is a natural fit within this project. First, Kriegel's view, like Husserl's, is a view according to which the Ubiquity Thesis is true. Having just critiqued Husserl's views on self-consciousness, it makes sense to examine another defence of the Ubiquity Thesis. Second, Kriegel's views on consciousness have a lot in common with Brentano's views. Husserl, you will recall, distanced himself from Brentano's ideas, some of which appear again in Kriegel's work, and so discussion of Kriegel's views on consciousness and self-consciousness provide a nice contrast to those of Husserl. Perhaps the most interesting contrast has to do with the fact that Kriegel argues that all conscious states represent themselves. Husserl explicitly argued against the Brentanian idea that all conscious states have themselves as intentional objects. Third, like Rosenthal's theory of consciousness discussed in chapter 5, Kriegel's theory of consciousness is a higher-order theory of consciousness. In this chapter, I will present Kriegel's views on consciousness and self-consciousness and then argue that there are problems with his account of self-consciousness.

7.1 Kriegel's Account of Subjective Character

Kriegel (2003a,b, 2004) has defended the Ubiquity Thesis in several works, most recently (2009) as part of his higher-order theory of consciousness. One of the significant features of Kriegel's view is that the Ubiquity Thesis *falls out* of his theory of consciousness. Although the Ubiquity Thesis falls out of his theory, Kriegel does not merely accept the thesis because it is a consequence of the theory. Rather, Kriegel maintains that the fact that his theory accommodates the Ubiquity Thesis is a reason to *prefer* his theory over other higher-order theories of consciousness. In this section, I will present a sketch of Kriegel's theory of consciousness, and say a few words about how he motivates the Ubiquity Thesis.

All conscious states have phenomenal character. Having phenomenal character is what distinguishes conscious from unconscious states. Phenomenal character, according to Kriegel (2009), can be broken down into two distinguishable components: *qualitative character* and *subjective character*. A conscious state's qualitative character is what makes it the *kind* of conscious state it is, and the qualitative character of a conscious state is determined by its representational content. When, for instance, I look at my coffee mug, the mental state I am in represents my coffee mug as being a certain way. And it is in virtue of the way my mental state represents my coffee mug that I am in a mental state with a reddish qualitative character. A conscious state's subjective character is what makes it a conscious state *at all*. According to Kriegel, a mental state has subjective character in virtue of the fact that it represents itself.

Kriegel motivates his theory of consciousness in the following way. A conscious state, he argues, is one that its subject must be aware of. This, he says, "is almost a conceptual truth" (2004, 198). "Conscious experiences are not states that we may *host*, as it were, *unawares*" (2009, 8). What must be the case in order for an individual to be aware of his conscious state? It must be represented. Awareness requires representation, so an individual cannot be aware of something unless that individual

is in a mental state that represents it. However, it is not enough that a conscious state be represented. In order for an individual to be aware of a conscious state, it must be represented in the *right sort of way*: it must be represented by a mental state that is itself a conscious state. The reason for this is that subjects of conscious states have “first-hand, first-person knowledge of [their] awareness of [their] conscious states” (2009, 19). If a subject’s conscious state was represented *unconsciously*, then he would not have the right sort of knowledge of his awareness of his conscious state. Thus, in order for a mental state to be a conscious mental state, it must itself be represented by a conscious mental state.

If Kriegel were defending a two-state higher-order theory of consciousness, this commitment would be problematic.¹ If every conscious mental state must itself be represented by a conscious mental state, and if the representing state must be a higher-order mental state that is, as many higher-order theories maintain, *numerically distinct* from the represented state, then if a subject has a single conscious state, he has infinitely many of them. Kriegel avoids this problem by claiming that the representing state and the represented state are identical. According to Kriegel’s theory, a mental state is conscious in virtue of its representing itself. And since the state that is doing the representing is a conscious state, its subject can have the appropriate knowledge of its awareness of the appropriate mental state. Since every conscious mental state represents itself, and, in virtue of representing itself, has subjective character, every conscious mental state has subjective character. This, in essence, is the Ubiquity Thesis. Kriegel’s precisification of it, which he calls the “Ubiquity of Inner Awareness Thesis,” is the following:

¹See section 5.1.

For any mental state M of a subject S , if M is conscious at t , then (i) S is aware of M at t and (ii) S 's awareness of M is part of S 's overall phenomenology at t . (2009, 181)

Kriegel anticipates the skeptic who doubts that one is always aware of one's conscious mental state. He responds by motivating a view according to which our inner awareness of our conscious states, though not focal, is always *peripheral*; we are always peripherally aware of our conscious states. To motivate this view, Kriegel considers some general features of outer awareness, i.e., our awareness of external objects, and then applies them to our inner awareness of our own mental lives. While I type, my eyes follow the cursor on the screen and I am visually aware of it. I am also visually aware of the book that is sitting on the table beside my computer. However, there is a difference between my visual awareness of the cursor on the screen and my visual awareness of the book on the table beside my computer. One might say that I am focally aware of the cursor and peripherally aware of the book. Similar distinctions can be made with respect to hearing. If I pause my typing for a moment and direct my attention to the discussion that is happening in the seminar room down the hall, then the discussion is the focus of my aural awareness. However, I am also peripherally aware of the sound of the fan running in my computer.

Lest anyone doubt that there is such a thing as peripheral outer awareness, Kriegel argues that there are functional reasons for countenancing peripheral outer awareness. Suppose there are five different objects in one's visual field. There are many different ways one's attentional resources could be allocated among those five objects. Kriegel (2009, 256) points out that, as a matter of contingent fact, about 60% of one's attentional resources are allocated to one of the objects, whichever happens to be the focus of one's attention, and about 10% of one's attentional resources are allocated to each of the other four objects. Kriegel calls this the "60/10 strategy." But things could have been different. Our visual and cognitive systems could have allocated our

attentional resources according to a 100/0 strategy, according to a 20/20 strategy, or any of a large number of other strategies. Had our visual and cognitive systems employed a 100/0 strategy, there would have been such a thing as focal visual awareness, but no such thing as peripheral visual awareness. Had our visual and cognitive systems employed a 20/20 strategy, there would have been no such thing as focal visual awareness, since no object in one's visual field would have received more attentional resources than any other.

Why, one might ask, do our visual and cognitive systems employ something like a 60/10 strategy instead of a significantly different strategy? Kriegel's suggestion is that there is a functional advantage to doing so. Our visual and cognitive systems have only limited resources at their disposal. They must "decide" which objects warrant a larger share of attentional resources and which objects require only a lesser share of attentional resources. Allocating visual attentional resources according to the 60/10 strategy has obvious survival benefits. Allocating a larger share of attentional resources at the object that is the focus of one's attention will make much more information about that object available to downstream processing systems, while allocating a smaller share of attentional resources to a few other objects still permits downstream processing systems to monitor them. In order to make the most of limited resources, our visual and cognitive systems must allocate attentional resources in certain ways, and the ways of allocating attentional resources that seem to provide the best chance of survival make us peripherally aware of certain objects in our visual fields.

Just as awareness can be outward-directed and perceptual, it can also be inward-directed and non-perceptual. One can, for instance, be aware of one's occurrent thoughts. And just as the distinction between focal and peripheral awareness can be applied to perception, it can also be applied to non-perceptual inner awareness. "Thus, as I focally think of tomorrow's lecture, I may also be peripherally aware

of the fact that I have yet to pay last month's electricity bill" (Kriegel, 2004, 191). This distinction between focal and peripheral awareness can also be applied across modalities. The example Kriegel gives is that of a long-distance truck driver who contemplates the meaning of life as she drives across the state of Nebraska. The focus of her awareness is on her inner thoughts. However, she remains peripherally aware of the cornfields that pass endlessly by.

This natural application of the distinction between focal and peripheral awareness to cases of non-perceptual awareness, argues Kriegel, provides us with an intuitive way to understand the distinction between reflective or transitive self-consciousness and minimal or, as Kriegel calls it, intransitive self-consciousness. When I attend to my internal state, as I do when I introspect and reflect on what it is like to be sitting at my computer, I am focally aware of my internal state and I am reflectively or transitively self-conscious. When I resume the process of writing, the focus of my attention shifts away from my internal state. I am no longer focally aware of it. However, even when my attention is focused outward, I am always peripherally aware of my conscious state. And since my peripheral awareness of my conscious state is part of my conscious state's phenomenology, my conscious state will have subjective character. Thus, in virtue of being in a conscious mental state, I will also be self-conscious.

7.2 Objections to Kriegel's Account

In this section, I will raise three different lines of criticism against Kriegel's account. First, I will quibble with some of the terminology with which Kriegel sets his views. Second, I will press Kriegel on his claim that peripheral awareness of our conscious states is phenomenologically manifest. Third, I will argue that, even if it is granted to Kriegel that we are *sometimes* aware of our conscious mental states, he

has not given us reason to believe that we are *always* aware of our conscious mental states.

7.2.1 Is Intransitive Self-Consciousness Really Intransitive?

The first criticism I will raise for Kriegel's view has to do with the way in which he sets out the conceptual terrain. In laying out his theory, Kriegel makes several distinctions. He distinguishes between focal and peripheral awareness, between perceptual and non-perceptual awareness, and between transitive and intransitive consciousness. The kind of self-consciousness that is ubiquitous, according to Kriegel, is intransitive self-consciousness. The distinction I want to call attention to is the distinction between transitive and intransitive consciousness and Kriegel's application of that distinction to self-consciousness.

Many have pointed out that the term "conscious" has both an intransitive and a transitive use.² Consider the following two statements:

Sam is conscious. (7.1)

Sam is conscious of the siren. (7.2)

In (7.1), "conscious" does not have an object and is used intransitively. The sentence appears to ascribe a non-relational property to Sam. In (7.2), the term has a direct object and is used transitively. The sentence appears to ascribe a relational property to Sam. The term can also be used to ascribe a property to a mental state. For example,

Sam's sensory state (involving the siren) is conscious. (7.3)

²See Armstrong and Malcolm (1984) for an early use of the distinction. Rosenthal (1997) is, perhaps, more well-known for the way in which he employed the distinction in his development of his higher-order theory of consciousness. Kriegel (2003a, 2004) also employs this distinction.

Kriegel (2004) calls the properties ascribed by “conscious” in (7.1) and (7.2) *intransitive creature consciousness* and *transitive creature consciousness* respectively, for they both ascribe a property to an individual. He calls the property ascribed by “conscious” in (7.3) *intransitive state consciousness*, for it ascribes a property to a mental state.³

One might wonder whether there is the further property of *transitive state consciousness*. In its transitive use, the term “conscious” only seems fit to ascribe a property to an individual. Sentences such as

Sam’s sensory state is conscious of the siren. (7.4)

in which the term is used transitively to ascribe a property to a mental state are awkward. In spite of the issues surrounding this transitive use of “conscious,” it is possible to delineate a property that one might call “transitive state consciousness.” For instance, even though the term “conscious” is most naturally used to ascribe properties to individuals, it is equally natural to think that the properties ascribed to individuals in such uses are derivative of the properties ascribed in (7.3) and (ignoring the awkwardness for the moment) in (7.4). If an individual has the property of intransitive creature consciousness, it is, plausibly, because of being in a mental state that has the property of intransitive state consciousness. Likewise, if an individual has the property of transitive creature consciousness, it is, plausibly, because that individual is in a state that bears a certain intentional relation to the object he or she is conscious of. We can say, as Kriegel does, that the state that bears the intentional relation to the object its subject is conscious of has the property of transitive state

³Like (7.1), (7.3) appears to ascribe a non-relational property. Higher-order theorists, however, will point out that the surface grammar of the sentence cannot be assumed to accurately report the metaphysical structure of reality. Rosenthal, for instance, maintains that “a state is intransitively conscious if one is transitively conscious of it” (1997, 738).

consciousness even if the transitive use of the term with respect to mental states is grammatically awkward. It should be noted, however, that, although it is possible to delineate the property of transitive state consciousness in this way, care must be taken to avoid a certain confusion. As Rosenthal says, “Being transitively conscious of something is a relation that a person or other creature bears to that thing. So only creatures can be transitively conscious of things. A mental state may well be that state in virtue of which somebody is conscious of a thing, but the state cannot itself literally be conscious of anything” (1997, 738).

Kriegel (2003a, 2004) extends the distinction between transitive and intransitive consciousness to self-consciousness. Substituting “self-conscious” for “conscious” in (7.1) and (7.2) respectively we get the following pair of sentences.

Sam is self-conscious. (7.5)

Sam is self-conscious of the siren. (7.6)

(7.5) contains an intransitive use of “self-conscious” and appears, *prima facie*, to ascribe a non-relational property, i.e., the property of *intransitive creature self-consciousness*, to Sam. (7.6) contains a transitive use of the term. If the parallelism between consciousness and self-consciousness obtains, then (7.6) ascribes the property of *transitive creature self-consciousness* to Sam.

There are several things to note about these sentences. First, (7.6) has an odd ring to it. Although it may be grammatically well-formed—something that is not entirely obvious—it is unclear how to understand it. There is nothing particularly mysterious about being conscious of a siren, but what does it mean to be *self*-conscious of a siren? To my ears, the most natural understanding of (7.6), if there is one, is that Sam is embarrassed by the siren. But this is not the understanding that Kriegel

advocates.⁴ He suggests that the sentence is equivalent to

Sam is conscious of his sensory state (involving the siren). (7.7)

(7.7) may be an appropriate way to understand (7.6), but, if so, it is not because (7.7) is clearly the most natural way to understand (7.6). In my opinion, the most natural way to understand a hyphenated expression like “self-conscious” is to understand “self” as indicating the object of consciousness, especially when we observe that “self” is often contrasted with “other.” The difference, for instance, between self- and other-consciousness would seem to be in what the subject of consciousness is conscious of. In light of this, (7.6) appears to be a clumsy attempt at ascribing two objects of consciousness to Sam: himself and the siren. However, if (7.7) is synonymous with (7.6), as Kriegel suggests, then neither he (Sam) nor the siren are asserted to be the object of Sam’s consciousness. Instead, Sam’s *sensory state* is the object.

Now, Sam’s sensory state could be thought of as being part of himself, and so being conscious of his sensory state would make him conscious of himself (i.e., self-conscious). This reveals that, according to Kriegel’s understanding of (7.6), it merely offers more information than (7.5); it specifies how Sam is self-conscious. All (7.6) does is indicate which part of himself Sam is conscious of. Here is the second point to make about sentences (7.5) and (7.6): as Kriegel himself observes, it would appear that the properties of intransitive and transitive creature self-consciousness are both dependent upon the same relational property of mental states. In the case of intransitive and transitive creature consciousness, the two properties were dependent upon different properties of mental states: intransitive and transitive state consciousness respectively. In the case of intransitive and transitive creature self-consciousness,

⁴I should mention that the sentences I have been using are not the same sentences Kriegel considers. I have attempted to ensure, however, that they have the same form.

however, the two properties are dependent upon the same property of mental states; the property of intransitive creature self-consciousness would seem to depend upon being in a mental state that has its subject for its object, and the same would seem to be true about the property of transitive creature self-consciousness. Because intransitive and transitive creature self-consciousness appear to be dependent upon the same property, this is at least one way in which the parallelism between intransitive and transitive consciousness and intransitive and transitive self-consciousness does not hold.

The fact that the parallelism between consciousness and self-consciousness is not perfect does not qualify as a problem for Kriegel's view. We can also grant Kriegel his reading of (7.6). What does count as a problem for Kriegel's view, however, is that the distinction between intransitive and transitive creature self-consciousness seems to be a distinction without a difference. As we have seen, both properties appear to be dependent upon the same relational property of mental states. Thus, there is no deep metaphysical difference between them.

One might point out that this merely places Kriegel in the same position as other higher-order theorists with respect to the way in which they understand intransitive and transitive creature consciousness. Higher-order theorists argue that an individual has the property of intransitive creature consciousness in virtue of being in a state that has the property of intransitive state consciousness. And intransitive state consciousness, they argue, depends upon transitive state consciousness. Thus, higher-order theorists seem committed to the claim that both intransitive and transitive creature consciousness depend upon transitive state consciousness.

Upon a closer look, however, there is a significant difference between Kriegel's distinction and the higher-order theorists' distinction between intransitive and transitive creature consciousness. Contrary to Kriegel's distinction between intransitive and transitive creature self-consciousness, the traditional distinction between intransitive

and transitive creature consciousness does rest upon a real metaphysical difference. Intransitive state consciousness, upon which the property of intransitive creature consciousness depends, is a different property than transitive state consciousness. A state that instantiates the latter is a state that carries intentional content about a state that instantiates the former, but states that instantiate the former do not typically carry intentional content about a state that instantiates the latter. Intransitive and transitive state consciousness are not identical properties, and so there is a real metaphysical difference between intransitive and transitive creature consciousness that does not obtain between intransitive and transitive creature self-consciousness.

This absence of a real metaphysical difference is compounded by the “pseudo intransitivity” of sentences that ascribe the property of intransitive creature self-consciousness. Hyphenated terms function grammatically as single terms, and because the hyphenated term “self-conscious” in (7.5) does not take an object, it is grammatically intransitive. However, as I have argued above, the most natural way to understand the hyphenated term is as providing the object of consciousness. Thus, although the hyphenated term is used intransitively in (7.5), it has a transitive sense that is built into it and difficult to ignore. One might argue, again, that in this regard Kriegel is no worse off than other higher-order theorists. After all, higher-order theorists maintain that the grammatical intransitivity of a sentence like (7.1) does not entail that consciousness is, at bottom, a non-relational property of mental states. But this response misses the point. Even if the property of creature consciousness is to be explained, ultimately, by transitive *state* consciousness, the fact remains that the property of creature consciousness is an intransitive property of *creatures*. My point about the pseudo-intransitivity of creature self-consciousness is that, whatever the underlying metaphysics happens to be, the property of creature self-consciousness is not, *pace* the grammatical status of (7.5), genuinely intransitive. Given that there is nothing particularly intransitive about intransitive creature

self-consciousness, it would seem that the property has been mis-named. A better name would be the one Kriegel uses for a putatively different property, “transitive creature self-consciousness,” but even this would be less than ideal. The term “transitive” in this expression is misleading and redundant. It is misleading in that it implies that the property has an intransitive correlate. It does not. There is only one self-consciousness property of creatures. It is redundant in that the hyphenated term “self-consciousness” already implies transitivity. Thus, the label “creature self-consciousness” is already adequate.

What about intransitive and transitive *state* self-consciousness? Kriegel defines *transitive state self-consciousness* as the property a subject’s mental state has in virtue of which that subject has the property of transitive creature self-consciousness. As was indicated above in our discussion of (7.6) and (7.7), this is a relational property; a mental state instantiates this property when it bears an intentional relation to another of its subject’s mental states. What is significant about the property of transitive state self-consciousness is that it is just the property of transitive state consciousness with a restriction on what the relatum of the intentional relation it stands in is; the property of transitive state consciousness is indifferent about what its instantiating state bears an intentional relation to, whereas the property of transitive state self-consciousness requires that its instantiating state bear an intentional relation to a mental state that has the same subject as it does.

There is nothing problematic about the property of transitive state self-consciousness, though, again, one might quibble with the name. “Transitive state self-consciousness” appears to be redundant in the same way that “transitive creature self-consciousness” is, and the name also implies that there is such a thing as intransitive state self-consciousness. I will have more to say about this below. Lastly, the name implies that mental states can be self-conscious (i.e., conscious of themselves). As we have seen, mental states are not, properly speaking, conscious of anything.

According to Kriegel, the property of *intransitive state self-consciousness* is the property that is ascribed to Sam's mental state in the following sentence:

Sam's sensory state (involving the siren) is self-conscious. (7.8)

This sentence appears to ascribe a property to a mental state that mental states cannot have. As was just noted, mental states cannot, properly speaking, be conscious of anything, including themselves. What, then, is the nature of the property that (7.8) ascribes to mental states? Kriegel suggests that a sentence like (7.8) is equivalent to

Sam is self-consciously sensing (the siren). (7.9)

A *prima facie* problem with this equivalency claim is that (7.8) ascribes a property to a mental state, whereas (7.9) ascribes a property to Sam. Perhaps this can be overlooked. Kriegel argues that, in (7.9), "self-consciously" modifies the term that follows it and so the sentence suggests that Sam is sensing in a particular way. Perhaps this can be construed as an indirect way of saying something about Sam's mental state, and so the problem with the equivalency claim may be nothing more than a *prima facie* problem.⁵ The bigger question is whether the property of intransitive state self-consciousness is genuinely different than the property of transitive state self-consciousness, and, if there is a difference, is the difference one that tracks the difference between intransitivity and transitivity?

According to Kriegel, the distinction between transitive self-consciousness and intransitive self-consciousness is supposed to track the distinction between focal awareness of one's mental states and peripheral awareness of one's mental states; tran-

⁵It should be noted, however, that although it *could* be construed in this way, nothing about the structure of the sentence *requires* that it be construed in this way. "Self-consciously" says something about the way in which Sam is sensing the siren, but this might involve completely different mental states than the sensation of the siren itself.

sitive self-consciousness is focal awareness of one's mental states, and intransitive self-consciousness is peripheral awareness of one's mental states. This means that, regardless of whether one is in a state of transitive self-consciousness or intransitive self-consciousness, one is in a mental state that has an object (i.e., one is in a transitive state). Moreover, the potential objects of the respective mental states overlap exactly; every potential object of a transitively self-conscious mental state is also a potential object of an intransitively self-conscious mental state and vice versa. To be sure, there is a difference between focal and peripheral awareness, but the difference has nothing to do with the fact that one type of mental state is a relational or transitive state and the other an intransitive or non-relational state. *Both* are relational, transitive states. Thus, to say that the one sort of state is intransitively self-conscious is misleading.

It should be pointed out that, although Kriegel employs the terminology of intransitive and transitive self-consciousness in earlier works (2003a, 2004), it is largely absent from his most recent work (2009). I am not aware that he explicitly distances himself from it, but perhaps one should infer from its absence in his later work that he no longer endorses it. If this is the case, the criticisms I have raised should be understood as being directed at the views presented in his earlier works and not his later ones.

I should also acknowledge that "intransitive self-consciousness" is a term of art for Kriegel. As such, any criticism of the expression or the way in which it is used will be of limited significance, and I certainly do not take the criticism I have just raised to be a decisive objection to Kriegel's view. This being said, defenders of the Ubiquity Thesis, like Kriegel, have certain dialectical goals, and their chances of achieving their goals depend somewhat upon how innocuous they make their views seem. If a defender of the Ubiquity Thesis says that a certain kind of self-consciousness is intransitive and explicitly draws some parallels between it and intransitive conscious-

ness, a notion already entrenched in the literature, the phenomenon is more likely to escape close scrutiny.

7.2.2 Is Peripheral Inner Awareness or Minimal Self-Consciousness Phenomenologically Manifest?

The list of philosophers who have accepted the Ubiquity Thesis is impressive. I have already registered my scepticism about the existence of pre-reflective or minimal self-consciousness. What motivates or grounds this scepticism? It is rooted in the fact that I fail to recognize it in my own experience; I fail to detect its phenomenological presence. This is not the case with respect to the more robust sort of self-consciousness mentioned in section 6.1; I believe I can isolate or “point to” some features of my experience that can plausibly be said to be *its* phenomenological footprint or presence. However, with respect to minimal self-consciousness, I do not have similar success.

Phenomenological disputes are common among philosophers of consciousness; one philosopher makes a claim about something he finds in his experience while another claims that she fails to find the same thing in her own experience. Claims about what one finds or does not find in one’s experience seldom, if ever, decide substantive disputes in the philosophy of consciousness, and I do not offer my examination of my own experience as a decisive objection to the existence of minimal or thin self-consciousness or, by extension, the Ubiquity Thesis. Nevertheless, I doubt I am the only one who fails to locate the phenomenological presence of self-consciousness. This means that, for many people, the existence of minimal or thin self-consciousness is called into question. Something must be said that will either help those of us who fail to recognize it in our experience recognize it or will bolster the Ubiquity Thesis against such scepticism. The discussion in this section will be anchored around Kriegel’s views, but my criticisms are also intended to have a larger scope.

To the inquirer who confesses failure in his attempts to locate minimal self-consciousness in his experience and yet remains open to the possibility of being educated as to how to go about finding it, Kriegel and others offer relatively little by way of help. Zahavi (2005, 11), for instance, characterizes minimal self-consciousness as “tacit.” Kriegel (2004, 191) says peripheral inner awareness is “subtle,” “peripheral,” “implicit,” and seems to endorse David Chalmers’ (1996, 10) characterization of it as “deep and intangible,” a “background hum.” These characterizations drive home the point that one should not expect the phenomenological manifestation of self-consciousness to announce itself loudly, but they do not offer much, if anything, by way of help to the open-minded inquirer that might help him find the phenomenological footprint of minimal self-consciousness if he has not already identified it.

Kriegel offers some explanation for why peripheral inner awareness is so difficult to locate phenomenologically. First, unlike focal inner awareness, peripheral inner awareness is not the product of introspection or reflection. This is so by definition. As soon as one attempts to introspect or focus upon it, the phenomenon disappears. If introspection or reflection is successful, peripheral inner awareness automatically gives way to focal inner awareness. Second, peripheral inner awareness is phenomenologically unimpressive. It runs quietly in the background. Third, peripheral inner awareness is ubiquitous. Focal inner awareness is much more rare. Thus, we can locate its phenomenological footprint by contrasting our experiences when we are in a state of focal inner awareness with our experiences when we are not in such a state. Fourth, peripheral inner awareness is involuntary, whereas focal inner awareness is voluntary. We can direct our attention inward and bring about a state of focal inner awareness. Peripheral inner awareness is not something we as agents have control over. These features of peripheral inner awareness all conspire to make it particularly difficult to locate or isolate its phenomenological footprint.

A defender of the Ubiquity Thesis could go even further. He could point out that, if the thesis is true, its defender is saddled with a certain methodological handicap. Suppose there was a person who claimed to be ignorant of what the phenomenological manifestation of a certain wavelength of visible light was, say light in the reddish range. One could help this person by bringing him within visual range of a stop sign and pointing to it. Similarly, if someone claimed not to know what it is like to perform certain mental exercises like, say, doing relatively simple arithmetic, there would be a straightforward way to enlighten him. One could, again, direct the person to attend to something in her experience. One could say, “Introspect on what it is like for you right now. Now add 25 and 15 in your head. What was it like for you while you were doing the calculation? Was it different for you while you were doing the calculation than it was for you before you were doing the calculation? If so, that is the phenomenological manifestation of that kind of mental exercise.” Even if this person was a hardened sceptic with respect to the existence of a distinctive phenomenology for certain mental exercises and this person claimed not to notice anything phenomenologically distinctive about the mental exercise, this would be the way to go about trying to convince her otherwise. However, when it comes to the person who claims ignorance with respect to the phenomenological manifestation of minimal self-consciousness, no analogous method for enlightening such an individual exists. The target hypothesis even implies as much. It claims that one is in a state of minimal self-consciousness whenever one is in a conscious state. Thus, if it is true, the method of contrast employed in the other cases will be unavailable to help enlighten the person who claims to be ignorant of what it is like to be in a state of minimal self-consciousness. After all, the hypothesis predicts, one will never be both conscious and not in a state of minimal self-consciousness, making it impossible to contrast what it is like to be in a state of minimal self-consciousness and what it is

like when one is not in a state of minimal self-consciousness. For when one is not in a state of minimal self-consciousness, one is not even conscious.

What these explanations show, I think, is that dismissal of the Ubiquity Thesis merely because of failure to locate the phenomenological footprint of minimal self-consciousness in one's experience would be, perhaps, too hasty. However, such failure remains a significant problem for the defender of the Ubiquity Thesis. The phenomenological manifestation of minimal self-consciousness may be subtle, involuntary, and ubiquitous, but, if it is a real phenomenon, there must be *something* that can be said about what it contributes to the what-it's-like-ness of experience. There must be something that the defender of the thesis can say to help the open-minded inquirer locate the relevant features of experience that are due to minimal self-consciousness. Failing this, doubt about the Ubiquity Thesis would seem to be warranted.

7.2.3 Are We *Always* in a State of Peripheral Inner Awareness?

In the previous section, I raised a skeptical challenge to the claim that peripheral inner awareness is phenomenologically manifest. In this section, I will consider a certain line of argument Kriegel gives for the Ubiquity Thesis. I will argue that even if it should turn out that we are sometimes in states of peripheral inner awareness, Kriegel has not given us good reason to think that we are always in states of peripheral inner awareness. Thus, he has not given us good reason to believe the Ubiquity Thesis.

One of the ways in which Kriegel motivates the claim that there is such a thing as peripheral inner awareness by considering our awareness of external objects.⁶ Kriegel points out that our awareness of external objects can be focal or peripheral. Given that this distinction applies in the case of our awareness of external objects, we have reason to believe that the same sort of distinction can be applied to intellectual

⁶See section 7.1.

awareness or our inner awareness of our own mental states. Kriegel says, “there appears to be peripheral awareness of every other sort, and it would be quite odd if the only exception was awareness of oneself” (2004, 192–193).⁷ Kriegel takes this to be an argument for the existence of peripheral inner awareness and, by extension, support for the Ubiquity Thesis. It is my contention that this sort of consideration provides little, if any, support for the Ubiquity Thesis.

Kriegel maintains that it would be odd if the distinction between focal and peripheral awareness applied only to outer awareness and not also to inner awareness. Although the tacit inference that is made in this claim is weak, the criticisms I will make of Kriegel’s view here will not challenge it. For the sake of the argument in this section, I will grant that there is such a thing as peripheral awareness of our own mental states. What I will argue, instead, is that even if it is possible for us to be peripherally aware of our own mental states, we do not yet have reason to accept the Ubiquity Thesis, which amounts to the claim, within the context of Kriegel’s theory, that we are always peripherally aware of our own mental states. My argument will proceed by considering, in some detail, a case of outer awareness and what it indicates about focal and peripheral outer awareness. Having done this, I will apply the results to inner awareness. The conclusion will be this: we have reason to believe that peripheral outer awareness of an object can fade into a complete lack of awareness of that object, so we have some reason to believe that peripheral inner awareness can likewise fade into complete unawareness.

As I sit at my desk in my office, typing at my computer keyboard, I hear a discussion in the seminar room down the hall. It distracts for a moment, but eventually I am able to draw my focus back to the task of writing my paper. Minutes pass. Eventually my writing reaches a bit of an impasse and, as I struggle to form my thoughts,

⁷See also Kriegel 2009, 50–51.

I am again distracted by the discussion happening down the hall in the seminar room. I have no recollection of what was said in the seminar room in the previous minutes. Indeed, no amount of introspection or reflection I can do at the moment gives me any reason to believe that I was even aware of anything that transpired in the seminar room over the last minutes. I do not have any memory of anything that was said. I do not have any recollection of there being even a muted murmur of voices in the background of my conscious experience over that period of time. As far as I can tell from my reflection upon my experience, I was completely unaware of what was going on in the seminar room while I was busy typing.

Now the mere fact that I have no introspective evidence that I was peripherally aware of what was going on in the seminar room while I was typing at my computer is not enough to show that I was not peripherally aware of what was going on in the seminar room. Perhaps I *was* peripherally aware but nothing I was peripherally aware of made it into long or short term memory. Given the nature of peripheral and focal awareness, there is not really any way to introspectively test whether I was peripherally aware while I was typing. I can, for instance, go back to typing and try to monitor my experience while I type to see if I am peripherally aware of the discussion down the hall. But if I do so, whenever I check to see whether I am peripherally aware, I become focally aware of the discussion in the seminar room. This is just the way focal awareness works.⁸ But given the lack of introspective evidence of peripheral awareness and the fact that one cannot test it, the possibility remains that I was not peripherally aware of the discussion in the seminar room at all.

Some will enter the following sort of counterfactual as evidence that I was peripherally aware of the discussion while I was typing. They will point out that, if someone in the seminar room had mentioned my name in the course of the discus-

⁸Kriegel (2009, 182–183) accepts an attention-shift model of introspection. Thus, for the reasons just outlined, Kriegel claims that it is *impossible* to introspect inner awareness.

sion, I would have picked up on it and my focal awareness would have been directed to the discussion. The counterfactual is probably true. But does it show that I was peripherally aware of the discussion? What it indicates, at very least, is that some sort of aural monitoring was going on. Perhaps this is enough to say that I was peripherally aware of the discussion. But even if this is conceded the question remains whether my awareness was *phenomenologically manifest*? The lack of introspective and reflective evidence of this would seem to indicate that it was not. Again, the mere lack of such evidence could not decide the question conclusively, but I think the lack of evidence clearly favors the hypothesis that it was not phenomenologically manifest over the hypothesis that it was. This is important, because, if it is possible to be completely unaware of external stimuli or if it is possible to be peripherally aware of external objects without that awareness being phenomenologically manifest, then merely maintaining that the distinction between focal and peripheral awareness applies equally to both outer and inner awareness will not be enough to show that individuals are in a state of peripheral inner awareness whenever they are conscious.

It is worth pointing out, here, that the above case seems to be importantly different from certain other cases that are mentioned by Kriegel in his discussion of focal and peripheral awareness. One example of peripheral awareness Kriegel mentions involves a long-distance truck driver who contemplates the meaning of life while driving across Nebraska. The driver's attention is focused inward as she thinks about the direction her life is going and the meaning of life, and she drives through miles of cornfields without really paying them any attention. Kriegel says that her awareness of the cornfields and the road ahead is peripheral. In this case, it is not very plausible to suppose that her peripheral awareness is not phenomenologically manifest. Perhaps this is due, in part, to the fact that the sensory modality of her peripheral awareness is the visual modality—it is difficult to imagine any sort of visual awareness, peripheral or focal, that does not have a phenomenological manifestation—but I think there is

another reason why we are not at all inclined to say that the trucker's peripheral awareness has no phenomenological footprint: the trucker's peripheral awareness is actively involved in the execution of a certain task. The truck driver must be able to keep her truck on the road, otherwise she and others might perish, and her awareness of the road is required for this. Her peripheral awareness can never fade away into a complete lack of awareness or terrible things will happen. My peripheral awareness of the discussion in the seminar room down the hall is not like the truck driver's peripheral awareness of the road at all. It is not a visual awareness and it is not required for any task I am doing. My peripheral awareness of the discussion can fade into a complete lack of awareness without compromising my functioning or my survival.

Kriegel's account of self-consciousness is built on the idea that inner awareness is like outer awareness in a certain respect: depending on how our attentional resources are allocated, we can be focally or merely peripherally aware of our internal states just like we can be focally or merely peripherally aware of external objects. If this is the model we are to use for thinking about inner awareness, however, we need to know exactly how far the similarity goes. If I examine my own experience, it seems to me as though my awareness of external objects in my immediate environment is constantly fluctuating from focal awareness to peripheral awareness and then to a complete lack of awareness, depending upon how my attentional resources are being allocated. If it is possible for my outer awareness of an object to fade to peripheral awareness and then out of awareness altogether (even when the object remains there to be perceived), should not the same be possible for the objects of inner awareness? And if it is possible for any given mental state to fade from awareness completely, is it not possible for my general awareness of my own states to fade completely?

Kriegel cannot accept an affirmative answer to the last question, though he can maintain a qualified affirmative response to the question before it. Nothing about

my understanding of inner awareness, he is likely to say, requires that we are always aware of *all* of our mental states; at any given time, a person is likely to be in many mental states he or she is not aware of at all. All Kriegel requires is that we are always (at least peripherally) aware of *some* of our mental states. In my opinion, however, analogizing outer and inner awareness in the way Kriegel does provides very little support for this claim.

Perhaps the strongest consideration the analogy provides for this claim comes from the fact that, while we are in a state of general wakefulness, we are always aware of at least some external objects. No matter where my attention is focused at any given moment, even if I am focusing as much of my attention as I can inwardly, it seems I am always at least peripherally aware of some of the objects in my immediate environment. If inner awareness is like outer awareness, one might be inclined to suppose that the converse might also be true: even if am directing as many attentional resources upon an external object as I can, I will always be at least peripherally aware of some of my internal states.

This line of reasoning is too quick. The assumption that one must always be at least peripherally aware of some external objects during times of wakefulness can be questioned. For instance, I find it quite plausible that experienced practitioners of various meditation techniques may well be able to focus their attention inwardly with enough concentration as to become completely oblivious, even if only for a short period, of their external environment and perhaps even their own bodies. If this is a real possibility, this would, in my view, go a long way to lend credence to the claim that awareness of external objects, even mere peripheral awareness, is not a necessary part of experience. Our attentional resources are most naturally directed outwardly. So if it is possible to combat the natural draw of our attentional resources in this way, how much more likely is it that focused attention on external objects might cause our awareness of our mental states to fade away completely?

I believe that this is made all the more plausible by the fact that the distribution of our attentional resources appears to be a zero-sum affair. Our attentional resources are limited and we cannot attend to an indefinite number of objects at a time. If we focus on one object (or a small number of objects), we do so to the exclusion of others. And when we change our focus to a different object (or set of objects) we sacrifice awareness of the things we were just attending to.

Kriegel, you will recall, motivated the claim that there is such a thing as peripheral outer awareness by pointing out that, for creatures with limited attentional resources like us, it is advantageous to allocate our attentional resources in certain ways. In the case of vision, it is advantageous to allocate a large portion of our attentional resources to the focal object in the visual field and much smaller portions of attentional resources to a few other objects. Although Kriegel's discussion centered on the allocation of visual attentional resources, one would expect that similar considerations apply with respect to other sensory modalities. The exact details might be different—the hearing system, for instance, might not allocate attentional resources in the same proportions or spread them out over as many objects—but one would expect there to be a similar rationing of attentional resources. Moreover, one would also expect there to be a system in place that monitors and allocates attentional resources on a more global level. This sort of system would allocate attentional resources across modalities. If engrossed in a novel, the system might allocate more of one's total attentional resources to the visual system. While sitting and listening to one's favourite piece of music, the system might take some of the attentional resources away from the visual system and allocate them to the hearing system. The global allocation would be still different in the middle of a yoga routine or concentrating on a logic problem.

The way in which our attentional resources are allocated matters. If our attentional resources are ever allocated to places where they are not necessary, we face the prospect of compromising our capacity to uptake information, of impairing our ability

to function, or, in extreme cases, jeopardizing our survival. Consider a prehistoric hunter whose family and companions depend upon him to have a successful hunt for their survival. He would be at a disadvantage if some of his attentional resources were directed inwardly at his mental states instead of being directed outwardly at the game he is tracking? The same would be true if, during the hunt, he himself became the hunted; if his attentional resources were directed inwardly at his mental states, he would be less likely to react successfully to an impending attack. Or consider the musician who is trying to execute a performance of a musical work or the athlete who must be in sync with her teammates while they run a play in attempt to score a goal. Both the musician's ability to perform the musical work and the athlete's ability to execute her team's play properly would be compromised if not all of their respective attentional resources were devoted to the task at hand. And there are examples of an even more mundane sort. One's reading comprehension is directly correlated with the focus one brings to the task. If any attentional resources are directed inwardly at one's mental states or at some external object, one's comprehension of the material suffers.

It must be acknowledged that the mere fact that it would be advantageous to creatures like us if certain facts were true about our attentional resources and the way they are allocated does not entail that they, in fact, are that way. That said, the biological realm is replete with examples of organisms that have many of the traits they need to flourish. This gives us some reason to believe that the mechanisms governing the way in which our attentional resources are allocated may well be calibrated so as to pull attentional resources away from objects that do not need to be monitored. Although we need to be aware of things in our surrounding environment, I submit that we are seldom in situations where we require that our own mental states be monitored or attended to.

None of the considerations I have raised in this section constitute a conclusive argument against Kriegel's claim that we are always peripherally aware of at least some of our mental states. That said, I do believe that they show us that Kriegel's view that inner awareness can be thought of as analogous to outer awareness is far from conclusive itself. If the Ubiquity Thesis is true, Kriegel has not provided us with much of a reason to think that it is.

CHAPTER 8

CONCLUSION

In this essay, I have examined two questions about the nature and structure of conscious experience: is conscious experience necessarily (phenomenally) unified and is consciousness necessarily accompanied by self-consciousness? I have argued that the answer to the former is affirmative and that the answer to the latter is negative.

The primary challenge against the Unity Thesis comes from empirical research on split-brain subjects. I argued against interpretations of the split-brain data that are inconsistent with the Unity Thesis and offered an interpretation of my own that is consistent with it. Two-stream interpretations of the data are able to offer a nice explanation of some of the disunified behaviours that split-brain subjects exhibit in experimental conditions, but they are not as well-situated to account for the unified behaviour split-brain subjects exhibit outside of the experimental conditions. The most serious cost for some two-stream interpretations, however, is that they seriously stretch the bounds of conceivability. Both the partial-unity account and the TSOS account face the challenge of saying how a single subject could have simultaneous experiences that do not have a conjoint phenomenology. I examined some ways of attempting to make sense of this possibility and argued that they all failed. The TSTS account avoids some of the conceivability worries, but it is the most difficult to reconcile with the unified behaviour of the split-brain subjects outside of the experimental conditions.

I also examined Bayne's Switch Model. The Switch Model, unlike the two-stream interpretations, is consistent with the Unity Thesis. I rejected the Switch Model,

however, because it relies on the problematic claim that split-brain subjects have only one conscious hemisphere at a time. On the face of it, the split-brain data can be seen to confirm the Switch Thesis. Upon closer examination, however, the data is not nearly as supportive. The account I favour is, like Bayne's Switch Model, consistent with the Unity Thesis, but, contra Bayne, maintains that both hemispheres of split-brain patients are conscious together. This saddles it with the particular challenge of accounting for how split-brain subjects can have a phenomenally unified consciousness while, at the same time, producing the kinds of disunified behaviours that split-brain subjects are famous for. I proposed a model according to which content becomes phenomenally conscious in virtue of being consumed by a system that is responsible for a motor or behavioural output; all content that is consumed by some system or other (at a time) is phenomenally conscious and part of a single maximal experience. This is how split-brain subjects have a phenomenally unified consciousness. The disunified behaviour, on the other hand, is accounted for by the fact that the different consuming systems, some of which may be relatively isolated to a single hemisphere, do not, as a result of the commissurotomy, all have equal access to sensory information. The system responsible for speech, for instance, is in the left hemisphere and only receives visual information from the right half of the visual field. Outside of the experimental conditions, these different systems are able to gain a more complete representation of reality and so their limitations with respect to how much sensory information they receive at a particular time are compensated for. This account, in essence, claims that split-brain subjects have a unified phenomenal consciousness and a disunified access-consciousness. Because this account can explain the split-brain data without contradicting the Unity Thesis, I believe it should be accepted over its rivals.

Rosenthal has offered an account of conscious unity that is an extension of his higher-order theory of consciousness. I considered two recent defences of higher-

order theories from the Mismatch Objection, Rosenthal's and Weisberg's, and argued that their responses fail. I then considered Rosenthal's account of conscious unity in light of the Mismatch Objection. The mental mechanisms that Rosenthal employs to account for conscious unity generate representational mismatches. Some of the representational mismatches generated by these mental mechanisms engage the Mismatch Objection. I concluded that, even though it may be possible to understand the gathering strategy in a way that avoids the Mismatch Objection, the common-ascription strategy is problematic.

In the last two chapters I took up the Ubiquity Thesis, the claim that an individual cannot be in a conscious state and fail to be self-conscious. Husserl, an early and influential theorist working in the phenomenological tradition, and Kriegel, a contemporary proponent of what could be considered a Brentanian theory of consciousness, both defend this thesis. I argued against the Ubiquity Thesis by raising objections against Husserl and Kriegel.

Husserl has a view according to which an object must be *given* to consciousness before it can become the primary object of a conscious state. This entails that, in order for consciousness to direct its intentional gaze upon itself in an act of reflective self-consciousness, consciousness must first be given to itself. This is commonly referred to as the *self-givenness* of consciousness. My primary objections to Husserl's view concerned this phenomenon. I argued, first, that consciousness is not given to itself when it is in the natural standpoint, that is, when its intentional focus is upon objects in the external environment. When a subject is engrossed in the affairs of day-to-day life, he or she lives through experience; what is given to a subject in the natural standpoint are the objects in the external environment and not one's own mental states. In order to engage in reflective self-consciousness, he or she must leave the natural standpoint, and it is only when this happens that consciousness is given to itself and it becomes possible to take his or her own mental states as primary objects

of consciousness. Whatever appeal the Ubiquity Thesis has is derived from the fact that it is easy to confuse background awareness of external objects with pre-reflective awareness of one's conscious states. I argued, second, that it is possible to understand inner time consciousness in a way that does not provide the structural footing for Husserl's claim that a pre-reflective self-consciousness pervades all of experience.

Kriegel defends a higher-order theory of consciousness according to which all conscious states represent themselves. The Ubiquity Thesis falls out of this theory. In Kriegel's terminology, all conscious states have the property of *intransitive state self-consciousness*. I quibbled with this terminology and then questioned Kriegel's identification of intransitive self-consciousness with peripheral inner awareness. Kriegel takes inner awareness to be analogous to outer awareness. I argued that there is no good reason to believe that we have a constant, peripheral awareness of our own mental states and that we have some reason to believe the contrary.

If my arguments in this essay are sound, I have gone some distance toward establishing that conscious experience is phenomenally unified but that it is not necessarily accompanied by self-consciousness.

BIBLIOGRAPHY

- Armstrong, David M. 1968. *A Materialist Theory of Mind*. London: Routledge.
- . 1981. “What is Consciousness?” In *The Nature of Mind and Other Essays*, 55–67. Ithaca, New York: Cornell University Press. Reprinted in Block et al. (1997).
- Armstrong, David M. and Malcolm, Norman. 1984. *Consciousness and Causality and Causality: A Debate on the Nature of Mind*. Great Debates in Philosophy Series. Oxford, England: Basil Blackwell.
- Baker, Lynne Rudder. 1998. “The First-Person Perspective: A Test for Naturalism.” *American Philosophical Quarterly* 35:327–348.
- . 2000. *Persons and Bodies: A Constitution View*. Cambridge, England: Cambridge University Press.
- Baumann, Peter. 2007. “Experiencing Things Together: What’s the Problem?” *Erkenntnis* 66:9–26.
- Bayne, Tim. 2008. “The Unity of Consciousness and the Split-Brain Syndrome.” *The Journal of Philosophy* 105:277–300.
- . 2010. *The Unity of Consciousness*. Oxford: Oxford University Press.
- Bayne, Tim and Chalmers, David. 2003. “What is the Unity of Consciousness?” In Axel Cleeremans (ed.), *The Unity of Consciousness: Binding, Integration, Dissociation*, 23–58. Oxford: Oxford University Press.

- Bisiach, E. and Luzzatti, G. 1978. "Unilateral Neglect of Representational Space." *Cortex* 14:129–133.
- Block, Ned. 1995. "On a Confusion about the Function of Consciousness." *Behavioral and Brain Sciences* 18:227–287. Reprinted in Block et al. (1997).
- . 2011. "The Higher Order Approach to Consciousness is Defunct." *Analysis* 71:419–431.
- Block, Ned, Flanagan, Owen, and Güzeldere, Güven (eds.). 1997. *The Nature of Consciousness: Philosophical Debates*. Cambridge, Massachusetts: The MIT Press.
- Bogen, Joseph E. and Gordon, Harold W. 1971. "Musical Tests for Functional Lateralization with Intracarotid Amobarbital." *Nature* 230:524–525.
- Butterworth, George. 1995. "The Self as an Object of Consciousness in Infancy." In Philippe Rochat (ed.), *The Self in Infancy: Theory and Research*, 35–51. Amsterdam: Elsevier Science B.V.
- . 1999. "A Developmental-Ecological Perspective on Strawson's 'The Self'." In Gallagher and Shear (1999), 203–211.
- Byrne, Alex. 1997. "Some Like It HOT: Consciousness and Higher-Order Thoughts." *Philosophical Studies* 86:103–129.
- Carnap, Rudolf. 1967. *The Logical Structure of the World*. Berkeley, California: University of California Press.
- Carruthers, Peter. 1996. *Language, Thought, and Consciousness: An Essay in Philosophical Psychology*. Cambridge, England: Cambridge University Press.
- . 2000. *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge, England: Cambridge University Press.

- Chabris, Christopher F., Weinberger, Adam, Fontaine, Matthew, and Simons, Daniel J. 2011. "You Do Not Talk About Fight Club If You Do Not Notice Fight Club: Inattentional Blindness For a Simulated Real-World Assault." *i-Perception* 2:150–153.
- Chalmers, David J. 1995. "Facing Up to the Problem of Consciousness." *Journal of Consciousness Studies* 2:200–219.
- . 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.
- Corballis, Michael C. 1995. "Visual Integration in the Split Brain." *Neuropsychologia* 33:937–959.
- Crick, Francis and Koch, Christof. 1995. "Are We Aware of Neural Activity in the Primary Visual Cortex." *Nature* 374:121–123.
- . 1998. "Consciousness and Neuroscience." *Cerebral Cortex* 8:97–107.
- Dainton, Barry. 2000. *Stream of Consciousness: Unity and Continuity in Conscious Experience*. London: Routledge.
- Dennett, Daniel C. 1978. "Where Am I?" In *Brainstorms: Philosophical Essays on Mind and Psychology*, 310–323. Montgomery, Vermont: Bradford Books.
- . 1991. *Consciousness Explained*. New York: Back Bay Books.
- Derrida, Jacques. 1973. *Speech and Phenomena*. Evanston, Illinois: Northwestern University Press. Translated by David B. Allison. Originally published in 1967.
- . 1990. *Le Problème de la genèse dans la philosophie de Husserl*. Paris: Presses Universitaires de France.
- Dretske, Fred. 1993. "Conscious Experience." *Mind* 102:263–283.

- Duval, Shelley and Wicklund, Robert A. 1972. *A Theory of Objective Self Awareness*. New York: Academic Press.
- Edelman, Gerald M. 1989. *The Remembered Present: A Biological Theory of Consciousness*. New York: Basic Books.
- Flanagan, Owen. 1992. *Consciousness Reconsidered*. Cambridge, Massachusetts: The MIT Press.
- Franz, Elizabeth A., Eliassen, James C., Ivry, Richard B., and Gazzaniga, Michael S. 1996. "Dissociation of Spatial and Temporal Coupling in the Bimanual Movements of Callosotomy Patients." *Psychological Science* 7:306–310.
- Gallagher, Shaun and Shear, Jonathan (eds.). 1999. *Models of the Self*. Thorverton, UK: Imprint Academic.
- Gazzaniga, Michael S. 2000. "Cerebral Specialization and Interhemispheric Communication: Does the Corpus Callosum Enable the Human Condition?" *Brain* 123:1293–1326.
- . 2005. "Forty-Five Years of Split-Brain Research and Still Going Strong." *Nature Reviews Neuroscience* 6:653–659.
- Gazzaniga, Michael S. and LeDoux, Joseph E. 1978. *The Integrated Mind*. New York: Plenum Press.
- Gazzaniga, Michael S. and Sperry, Roger W. 1966. "Simultaneous Double Discrimination Response Following Brain Bisection." *Psychonomic Science* 4:261–262.
- Goldman, Alvin I. 1970. *A Theory of Human Action*. New Jersey: Prentice-Hall Inc.
- Gopnik, Alison and Astington, Janet W. 1988. "Children's Understanding of Representational Change and its Relation to the Understanding." *Child Development* 59:26–37.

- Heidegger, Martin. 1982. *The Basic Problems of Phenomenology*. Bloomington, Indiana: Indiana University Press. Translated by Albert Hofstadter.
- Henry, Michel. 1973. *The Essence of Manifestation*. The Hague, Netherlands: Martinus Nihoff. Translated by Girard Etzkorn.
- Hume, David. 2000. *A Treatise of Human Nature*. Oxford: Oxford University Press.
- Hurley, Susan L. 1998. *Consciousness in Action*. Cambridge, Massachusetts: Harvard University Press.
- Husserl, Edmund. 1958. *Ideas: General Introduction to Pure Phenomenology*. New York: The Macmillan Company. Trans. W. R. Boyce Gibson. Originally published in 1913.
- . 1970. *Logical Investigations*, volume 2. London: Routledge & Kegan Paul. Trans. J. N. Findlay.
- . 1991. *On the Phenomenology of the Consciousness of Internal Time*. Dordrecht, The Netherlands: Kluwer Academic Publishers. Trans. John Barnett Brough. Originally published in 1928.
- Hyman Jr., Ira E., Boss, Matthew, Wise, Breanne M., McKenzie, Kira E., and Caggiano, Jenna M. 2010. “Did You See the Unicycling Clown? Inattentional Blindness While Walking and Talking on a Cell Phone.” *Applied Cognitive Psychology* 24:597–607.
- James, William. 1981. *The Principles of Psychology*. Cambridge, Massachusetts: Harvard University Press. Originally published in 1890.
- Kapitan, Tomis. 1999. “The Ubiquity of Self-Awareness.” *Grazer Philosophische Studien* 57:17–44.

- Koch, Christof. 2004. *The Quest for Consciousness: A Neurobiological Approach*. Englewood, Colorado: Roberts and Company Publishers.
- Kriegel, Uriah. 2003a. "Consciousness as Intransitive Self-Consciousness: Two Views and an Argument." *Canadian Journal of Philosophy* 33:103–132.
- . 2003b. "Consciousness as Sensory Quality and as Implicit Self-Awareness." *Phenomenology and the Cognitive Sciences* 2:1–26.
- . 2004. "Consciousness and Self-Consciousness." *Monist* 87:185–209.
- . 2009. *Subjective Consciousness: A Self-Representational Theory*. Oxford: Oxford University Press.
- Landis, Theodor, Assal, Gil, and Perret, Etienne. 1979. "Opposite Cerebral Hemispheric Superiorities for Visual Associative Processing of Emotional Facial Expressions and Objects." *Nature* 278:739–740.
- Levine, Joseph. 2001. *Purple Haze: The Puzzle of Consciousness*. Oxford: Oxford University Press.
- Levy, Jerre. 1969. *Information Processing and Higher Psychological Functions in Disconnected Hemispheres of Human Commissurotomy Patients*. Ph.D. thesis, California Institute of Technology. Unpublished.
- . 1990. "Regulation and Generation of Perception in the Asymmetric Brain." In Colwyn Trevarthen (ed.), *Brain Circuits and Functions of Mind: Essays in Honor of Roger W. Sperry*, 231–248. Cambridge, England: Cambridge University Press.
- Levy, Jerre and Trevarthen, Colwyn. 1976. "Metacontrol of Hemispheric Function in Human Split-Brain Patients." *Journal of Experimental Psychology: Human Perception and Performance* 2:299–312.

- Levy, Jerre, Trevarthen, Colwyn, and Sperry, Roger W. 1972. "Perception of Bilateral Chimeric Figures Following Hemispheric Deconnexion." *Brain* 95:61–78.
- Lewis, Michael. 1992. *Shame: The Exposed Self*. New York: The Free Press.
- . 2003. "The Development of Self-Consciousness." In Johannes Roessler and Naomi Eilan (eds.), *Agency and Self-Awareness: Issues in Philosophy and Psychology*, 275–295. Oxford: Oxford University Press.
- Lockwood, Michael. 1989. *Mind, Brain, and the Quantum: The Compound 'I'*. Cambridge, Massachusetts: Basil Blackwell.
- . 1994. "Issues of Unity and Objectivity." In Christopher Peacocke (ed.), *Objectivity, Simulation, and the Unity of Consciousness: Current Issues in the Philosophy of Mind*, volume 83 of *Proceedings of the British Academy*, 89–95. Oxford: Oxford University Press.
- Lycan, William. 1987. *Consciousness*. Cambridge, Massachusetts: The MIT Press.
- . 1996. *Consciousness and Experience*. Cambridge, Massachusetts: The MIT Press.
- Mark, Victor. 1996. "Conflicting Behavior in a Split-Brain Patient: Support for Dual Consciousness." In Stuart R. Hameroff, Alfred W. Kaszniak, and Alwyn C. Scott (eds.), *Toward a Science of Consciousness*, 189–196. Cambridge, Massachusetts: The MIT Press.
- Marks, Charles E. 1981. *Commissurotomy, Consciousness, and Unity of Mind*. Cambridge, Massachusetts: The MIT Press.
- Matey, Jennifer. 2006. "Two HOTs to Handle: The Concept of State Consciousness in the Higher-Order Thought Theory of Consciousness." *Philosophical Psychology* 19:151–175.

- Mathews, Marlon S., Linskey, Mark E., and Binder, Devin K. 2008. "William P. van Wagenen and the First Corpus Callosotomies for Epilepsy." *Journal of Neurosurgery* 108:608–613.
- Milner, A. D. and Dunne, J. J. 1977. "Lateralised Perception of Bilateral Chimaeric Faces by Normal Subjects." *Nature* 268:175–176.
- Nagel, Thomas. 1971. "Bisection and the Unity of Consciousness." *Synthese* 22:396–413.
- . 1974. "What Is It Like to Be a Bat?" *The Philosophical Review* 83:435–450.
- Neander, Karen. 1998. "The Division of Phenomenal Labor: A Problem for Representational Theories of Consciousness." *Philosophical Perspectives* 12:411–434.
- O'Doherty, Joseph E., Lebedev, Mikhail A., Ifft, Peter J., Zhuang, Katie Z., Shokur, Solaiman, Bleuler, Hannes, and Nicolelis, Miguel A. L. 2011. "Active Tactile Exploration Using a Brain-Machine-Brain Interface." *Nature* 479:228–231.
- Pais-Vieira, Miguel, Lebedev, Mikhail, Kunicki, Carolina, Wang, Jing, and Nicolelis, Miguel A. L. 2013. "A Brain-to-Brain Interface for Real-Time Sharing of Sensorimotor Information." *Scientific Reports* 3: Article No. 1319.
- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Puccetti, Roland. 1981. "The Case for Mental Duality: Evidence From Split-Brain Data and Other Considerations." *The Behavioral and Brain Sciences* 4:93–123.
- Rosenthal, David M. 1986. "Two Concepts of Consciousness." *Philosophical Studies* 49:329–359. Reprinted in Rosenthal (2005a).
- . 1993. "Multiple Drafts and Higher-Order Thoughts." *Philosophy and Phenomenological Research* 53:911–918.

- . 1997. “A Theory of Consciousness.” In Block et al. (1997), 729–753.
- . 2000. “Introspection and Self-Interpretation.” *Philosophical Topics* 28:201–233.
Reprinted in Rosenthal (2005a).
- . 2002. “Explaining Consciousness.” In David J. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*, 406–421. Oxford, England: Oxford University Press.
- . 2005a. *Consciousness and Mind*. Oxford: Oxford University Press.
- . 2005b. “Sensory Qualities, Consciousness, and Perception.” In Rosenthal (2005a), 175–226.
- . 2005c. “Unity of Consciousness and the Self.” In Rosenthal (2005a), 339–363.
- . 2011. “Exaggerated Reports: Reply to Block.” *Analysis* 71:431–437.
- Roskies, Adina L. 1999. “The Binding Problem.” *Neuron* 24:7–9.
- Sartre, Jean-Paul. 1956. *Being and Nothingness: An Essay on Phenomenological Ontology*. New York: The Philosophical Library. Translated by Hazel E. Barnes.
- . 1957. *The Transcendence of the Ego*. New York: The Noonday Press. Translated by Forrest Williams and Robert Kirkpatrick.
- Schechter, Elizabeth. 2012. “The Switch-Model of Split-Brain Consciousness.” *Philosophical Psychology* 25:203–226.
- Seager, William. 2004. “A Cold Look at the HOT Theory.” In Rocco J. Gennaro (ed.), *Higher-Order Theories of Consciousness: An Anthology*, 255–275. Philadelphia: John Benjamins Publishing Company.
- Searle, John R. 2000. “Consciousness.” *Annual Review of Neuroscience* 23:557–578.
Reprinted in Searle (2002).

- . 2002. *Consciousness and Language*. Cambridge, England: Cambridge University Press.
- Siewert, Charles. 2001. "Self-Knowledge and Phenomenal Unity." *Noûs* 35:542–568.
- Simons, Daniel J. and Chabris, Christopher F. 1999. "Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events." *Perception* 28:1059–1074.
- Simons, Daniel J. and Levin, Daniel T. 1998. "Failure to Detect Changes to People During a Real-World Interaction." *Psychonomic Bulletin & Review* 5:644–649.
- Smith, David Woodruff. 1989. *The Circle of Acquaintance: Perception, Consciousness, and Empathy*, volume 205 of *Synthese Library: Studies in Epistemology, Logic, Methodology, and Philosophy of Science*. Dordrecht: Kluwer Academic Publishers.
- Smith, David Woodruff and McIntyre, Ronald. 1982. *Husserl and Intentionality: A Study of Mind, Meaning, and Language*. Boston, Massachusetts: D. Reidel Publishing Company.
- Sperry, Roger W. 1968a. "Hemisphere Deconnection and Unity in Conscious Awareness." *American Psychologist* 23:723–733.
- . 1968b. "Mental Unity Following Surgical Disconnection of the Cerebral Hemispheres." In *The Harvey Lectures*, Series 62, 293–323. New York: Academic Press.
- . 1984. "Consciousness, Personal Identity and the Divided Brain." *Neuropsychologia* 22:661–673.
- Sperry, Roger W., Zaidel, Eran, and Zaidel, Dahlia W. 1979. "Self Recognition and Social Awareness in the Deconnected Minor Hemisphere." *Neuropsychologia* 17:153–166.

- Treisman, Anne. 1999. "Solutions to the Binding Problem: Progress Through Controversy and Convergence." *Neuron* 24:105–110.
- Trevarthen, Colwyn. 1962. "Double Visual Learning in Split-Brain Monkeys." *Science* 136:258–259.
- . 1970. "Experimental Evidence for a Brain-stem Contribution to Visual Perception in Man." *Brain, Behavior, and Evolution* 3:338–352.
- . 1974. "Functional Relations of Disconnected Hemispheres with the Brain Stem, and with Each Other: Monkey and Man." In Marcel Kinsbourne and Wallace Lynn Smith (eds.), *Hemispheric Disconnection and Cerebral Function*, 187–207. Springfield, Illinois: C. C. Thomas.
- Trevarthen, Colwyn and Sperry, Roger W. 1974. "Perceptual Unity of the Ambient Visual Field in Human Commissurotomy Patients." *Brain* 96:547–570.
- Tye, Michael. 2003. *Consciousness and Persons*. Cambridge, Massachusetts: The MIT Press.
- Weisberg, Josh. 2011. "Abusing the Notion of What-it's-like-ness: A Response to Block." *Analysis* 71:438–443.
- White, Stephen L. 1987. "What is it Like to be a Homunculus?" *Pacific Philosophical Quarterly* 68:148–174. Reprinted in White (1991).
- . 1991. *The Unity of Self*. Cambridge, Massachusetts: The MIT Press.
- Wilkes, Kathleen V. 1978. "Consciousness and Commissurotomy." *Philosophy* 53:185–199.
- Wimmer, Heinz and Perner, Josef. 1983. "Beliefs About Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception." *Cognition* 13:103–128.

- Zahavi, Dan. 1999. *Self-Awareness and Alterity: A Phenomenological Investigation*. Northwestern University Studies in Phenomenology and Existential Philosophy. Evanston, Illinois: Northwestern University Press.
- . 2000. “Self and Consciousness.” In Dan Zahavi (ed.), *Exploring the Self: Philosophical and Psychopathological Perspectives on Self-Experience*, volume 23 of *Advances in Consciousness Research*, 55–74. Philadelphia: John Benjamins Publishing Company.
- . 2002. “First-Person Thoughts and Embodied Self-Awareness: Some Reflections on the Relation Between Recent Analytical Philosophy and Phenomenology.” *Phenomenology and the Cognitive Sciences* 1:7–26.
- . 2003a. *Husserl’s Phenomenology*. Stanford, California: Stanford University Press.
- . 2003b. “Phenomenology of Self.” In Tilo Kircher and Anthony David (eds.), *The Self in Neuroscience and Psychiatry*, 56–75. Cambridge, England: Cambridge University Press.
- . 2004. “The Embodied Self-Awareness of the Infant: A Challenge to the Theory-Theory of Mind.” In Dan Zahavi, Thor Grünbaum, and Josef Parnas (eds.), *The Structure and Development of Self-Consciousness: Interdisciplinary Perspectives*, 35–63. Philadelphia: John Benjamins Publishing Company.
- . 2005. *Subjectivity and Selfhood: Investigating the First-Person Perspective*. Cambridge, Massachusetts: The MIT Press.
- . 2006. “Thinking about (Self-)Consciousness: Phenomenological Perspectives.” In Uriah Kriegel and Kenneth Williford (eds.), *Self-Representational Approaches to Consciousness*, 273–295. Cambridge, Massachusetts: The MIT Press.

Zahavi, Dan and Parnas, Josef. 1999. "Phenomenal Consciousness and Self-Awareness." In Gallagher and Shear (1999), 253–270.